THE PEDAGOGICAL SEMINARY AND

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY

CARL MURCHISON

John E. Anderson
University of Minnesota

Charlotte Bühler
Universität, Wien

William H. Burnham
Clark University

Cyril Burt
University of London

Leonard Carmichael
Tufts College

Ed. Claparède
Université de Genève

Edmund S. Conklin
Indiana University

Sante De Sanctis
R. Università di Roma

Arnold Gesell
Yale University

William Healy
Judge Baker Foundation,
Boston

Leta S. Hollingworth
Teachers College,
Columbia University

Buford Johnson
The Johns Hopkins University

Harold E. Jones
University of California

Truman L. Kelley
Harvard University

Yoshihide Kubo
Hiroshima Normal College

K. S. Lashley
Harvard University

A. R. Luria
Medico-biological Institute, Moscow

Toshio Nogami
Kyoto Imperial University

Henri Piéron
Université de Paris

Calvin P. Stone
Stanford University

Lewis M. Terman
Stanford University

Godfrey Thomson
University of Edinburgh

E. L. Thorndike
Teachers College,
Columbia University

C. J. Warden
Columbia University

John B. Watson
New York City

Helen Thompson Woolley
Teachers College,
Columbia University

If this space should be unstamped, this is the regular library edition
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not
be resold for a period of five years from date of publication

VOLUME 54

1939

Copyright, 1939, by The Journal Press

Published quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Abel, Theodora M 391
Ames, Louise Bates 313
Barker, Roger G 27
Beckham, Albert Sidney 327
Benton, Arthur L 455
Carmichael, Leonard 425
Cason, Hulsey 359
Dennis, Wayne 217
Forbes, Marie L H 235
Forlano, G. 353
Hagmann, Edward A 455
Higgons, Reginald A. 165
Horowitz, Eugene L 337
Householder, A S 249
James, W T 151
Kennedy, John L 119
Kinder, Elaine F. 107
Kumin, Edythe 265
Lane, Elizabeth B. 107
Lunger, Ruth 457
Mann, Cecil W. 435
Maudry, Maria 193
Meltzer, H 403
Meyer, Max F. 241
Mitchell, Mildred B 3
Moore, Clarence Carl 285
Moore, Herbert 481
Morris, Charles M. 85
Nekula, Maria 193
Nelson, Virginia Lafayette 181
Page, James D 457
Peatman, John Gray 165
Pintner, R 353
Richards, T W 181
Russell, R W 217
Ryans, David G 223, 229, 461, 467
Sill, Jane B 391
Smith, M F 425
Smith, Randolph B 337
Smith, Stevenson 17
Stone, Calvin P 27
Tinklepaugh, Otto L 3
Wheeler, Lester R 295
White, Mary A 73
Williams, Harold M 73
Wolf, Ernst 475
Wood, Louise 265

TABLE OF CONTENTS

Monthly and weekly weight cycles in women and their relations to behavioral and physiological functions 3
By Otto L. Tinklepaugh and Mildred B. Mitchell

Age and sex differences in children's opinion concerning sex differences 17
By Stevenson Smith

The attitudes and interests of premenarcheal and postmenarcheal girls 27
By Calvin P. Stone and Roger G. Barker

The approach-withdrawal pattern in the social behavior of young children 73
By Mary A. White and Harold M. Williams

A critical analysis of certain performance tests 85
By Charles M. Morris

Relativism in the thinking of subnormal subjects as measured by certain of Piaget's tests . 107
By Elizabeth B. Lane and Elaine F. Kinder

The effects of complete and partial occipital lobectomy upon thresholds of visual real movement discrimination in the cat 119
By John L. Kennedy

Further experiments in social behavior among dogs . . 151
By W. T. James

Height-weight variability from birth to five years of age for children reared with optimal pediatric and home care 165
By John Gray Pearman and Reginald A. Higgons, M.D.

Studies in mental development III Performance of twelve-months-old children on the Gesell schedule, and its predictive value for mental status at two and three years . . 181
By Virginia Lafayette Nelson and T. W. Richards

Social relations between children of the same age during the first two years of life . . . 193
By Maria Maudry and Maria Nekula

SHORT ARTICLES AND NOTES . . 217

Comments on recent studies of *VTE* (Wayne Dennis and R. W. Russell) 217

"Persistence" test scores of students compared with respect to the nativity of the male parent (David G. Ryans) . . 223

A tentative statement of the relation of persistence test scores to certain personality traits as measured by the Bernreuter inventory (David G. Ryans) . . . 229

APPARATUS 235

Grooved blocks for use in registering choice (Marie I. H. Forbes) 235

BOOKS 237

CRITICAL REVIEWS OF RECENT BOOKS 241

S. S. Stevens and H. Davis *Hearing: Its Psychology and Physiology.* (Reviewed by Max F. Meyer) . . 241

K. Lewin. *Principles of Topological Psychology* (Reviewed by A. S. Householder) 249

BOOKS RECENTLY RECEIVED 261

A new standardization of the Ferguson form boards . . 265
By Louise Wood and Edythe Kumin

The elementary school mark . . . 285
By Clarence Carl Moore

A comparative study of the difficulty of the 100 addition combinations 295
By Lester R. Wheeler

Some relationships between stair climbing and prone progression . 313
By Louise Bates Ames

The intelligence of a negro high school population in a northern city 327
By Albert Sidney Beckham

Social relations and personality patterning in preschool children . 337
By Eugene L. Horowitz and Randolph B. Smith

Season of birth and intelligence 353
By R. Pintner and G. Forlano

An intelligent-question method of teaching and testing . . . 359
By Hulsey Cason

The perceiving and thinking of normal and subnormal adolescents and children on a simple drawing task . . 391
By Theodora M. Abel and Jane B. Sill

Nationality preferences and stereotypes of colored children . 403
By H. Meltzer

Quantified pressure stimulation and the specificity and generality of response in fetal life 425
By Leonard Carmichael and M. F. Smith

A test of general ability in Fiji 435
By Cecil W. Mann

SHORT ARTICLES AND NOTES 455

Psychometric test results in two cases of precocious puberty (Arthur L. Benton and Edward A. Hagmann) . . . 455

Worries of college freshmen (Ruth Lunger and James D Page) 457

Changes in variability in "digit-symbol substitution" performance measured at the beginning and at the end of practice (D. G Ryans) 461

An observation of the changes in variability of high, medium, and low "intelligence" groups, measured at the beginning and at the end of "digit-symbol substitution" practice (D. G. Ryans) 467

BOOKS 471

CRITICAL REVIEWS OF RECENT BOOKS 475

B F Skinner *The Behavior of Organisms.* (Reviewed by Ernst Wolf) 475

C Spearman *Psychology Down the Ages* (Reviewed by Herbert Moore) 481

BOOKS RECENTLY RECEIVED 487

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out. Beginning with January, 1936, the following methods and terms will be in force:

(1) The size of each volume will be 400-600 pages, and the subscription price will be $7 per volume, two volumes per year.

(2) The format of each volume will be such that each subscriber may tear down the assembled volume into its component articles.

(3) Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4) The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half-volume units. The second unit of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated time.

(5) Each author will be charged the actual manufacturing costs of any zinc or half-tone cuts.

(6) The publication charges will be as follows:

4	page signature	$12
8	" "	16
12	" "	24
16	" "	32
20	" "	40
24	" "	48
28	" "	56
32	" "	64
36	" "	72
40	" "	80

These charges are tentative and will be adjusted as justified by the subscription list.

(7) The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, to accommodate publication needs.

Beginning with January, 1938, eight hundred to twelve hundred pages (two volumes) annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7 00 per volume
Single numbers $4 00

QUARTERLY
Two volumes per year

March, 1939
Volume 54, First Half

Founded by G Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND
JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication

MARCH, 1939

(Manuscripts are printed in the order of final acceptance)

Monthly and weekly weight cycles in women and their relations to behavioral and physiological functions 3
By Otto L Tinklepaugh and Mildred B Mitchell

Age and sex differences in children's opinion concerning sex differences 17
By Stevenson Smith

The attitudes and interests of premenarcheal and postmenarcheal girls 27
By Calvin P Stone and Roger G Barker

(OVER)

Copyright, 1939, by The Journal Press

Provincetown, Massachusetts
Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass, under the act of March 3, 1879
Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass, under the Act of March 3, 1879

The approach-withdrawal pattern in the social behavior of young children 73
By Mary A White and Harold M. Williams

A critical analysis of certain performance tests . . . 85
By Charles M. Morris

Relativism in the thinking of subnormal subjects as measured by certain of Piaget's tests 107
By Elizabeth B Lane and Elaine F. Kinder

The effects of complete and partial occipital lobectomy upon thresholds of visual real movement discrimination in the cat . . . 119
By John L. Kennedy

Further experiments in social behavior among dogs 151
By W. T James

Height-weight variability from birth to five years of age for children reared with optimal pediatric and home care 165
By John Gray Peatman and Reginald A Higgons, M D.

Studies in mental development: III. Performance of twelve-months-old children on the Gesell schedule, and its predictive value for mental status at two and three years 181
By Virginia Lafayette Nelson and T. W. Richards

Social relations between children of the same age during the first two years of life 193
By Maria Maudry and Maria Nekula

SHORT ARTICLES AND NOTES 217

Comments on recent studies of *VTE* (Wayne Dennis and R W. Russell) 217

"Persistence" test scores of students compared with respect to the nativity of the male parent (David G. Ryans) 223

A tentative statement of the relation of persistence test scores to certain personality traits as measured by the Bernreuter inventory (David G. Ryans) 229

APPARATUS 235

Grooved blocks for use in registering choice (Marie L H. Forbes) 235

BOOKS 237

CRITICAL REVIEWS OF RECENT BOOKS . . 241

S S Stevens and H Davis *Hearing, Its Psychology and Physiology* (Reviewed by Max F Meyer) . . . 241

K Lewin *Principles of Topological Psychology* (Reviewed by A S. Householder) 249

BOOKS RECENTLY RECEIVED 261

The Journal of Genetic Psychology, 1939, **54**, 3-16

MONTHLY AND WEEKLY WEIGHT CYCLES IN WOMEN AND THEIR RELATIONS TO BEHAVIORAL AND PHYSIOLOGICAL FUNCTIONS*

New York City and the Psychopathic Hospital of Iowa City

Otto L. Tinklepaugh and Mildred B. Mitchell

This study was undertaken to determine the presence or absence of periodic weight changes correlated with the menstrual cycle in women. The results, while verifying the presence of such monthly cycles, are perhaps more important because they bring to attention weight changes of a weekly periodicity which result from behavioral factors. The weekly cycles tend to obscure the monthly cycles, and they may reflect weekly changes in both physical condition and behavioral tendencies that distort data secured in many forms of psychological and physiological study.

Tinklepaugh (8) discovered a clearly defined weight cycle in a female chimpanzee and later verified his finding in an unpublished study of a group of these animals. Women not uncommonly believe that their weight is greater just prior to menstruation, though this seems never to have been verified experimentally

A. Source of Data

Data for the present study were secured by Dr. Mitchell while she was a member of the staff of George School, George School, Pa Two teachers, and 26 girl students ranging from 14 to 21 years in age, served as subjects. Two of the girls had menstruated only once or twice and were not yet experiencing regular menstrual cycles. The other members of the group menstruated at intervals of from 14 to 55 days, with a mode of 28 days and a mean of 31 5 days (Figure 2, No. 4)

The study was begun in late January and continued for 20 weeks. During that time, except for days of illness, occasional week-end absences, and a one-week vacation, the subjects lived in a dormi-

---

*Received in the Editorial Office on March 1, 1938.

tory where they ate at the same tables, followed similar scholastic and social programs, retired at similar hours, and in most other respects lived in like manner.

The subjects were weighed nightly after they had prepared for bed, usually at 10:00 P.M. The scales used were of the beam type, calibrated to one-fourth pound.

Records of illness, absence from school, and the onset and duration of the menstrual flow were kept.

B. Weekly Weight Cycles

When the daily weights of each of the subjects were plotted, the resulting curves were highly irregular and upon casual inspection most of them revealed little evidence of periodic changes. Menstrual

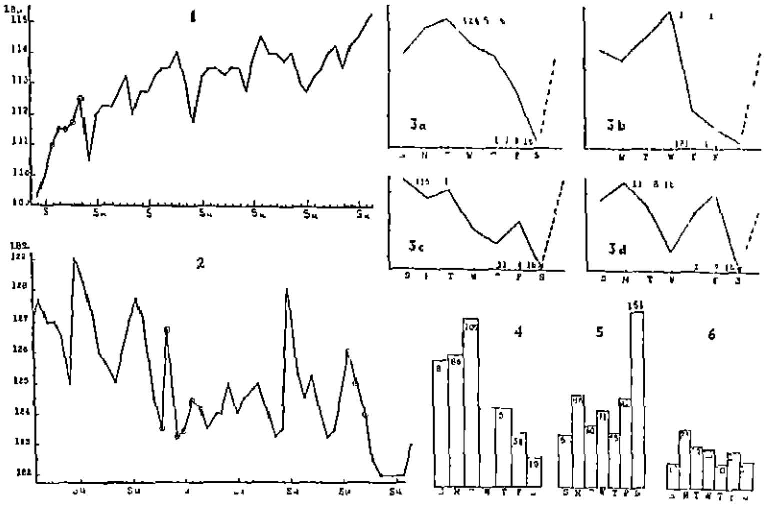


FIGURE 1

Weekly Weight Cycles

(1) Unusually regular cycles of Subject 19 when she was recovering from illness. Circles indicate days of menstruation (2) Moderately regular weekly cycles superimposed on monthly cycles of Subject 22 (*3a, b, c, d*) Typical average weekly fluctuations of individual subjects over 20 weeks. Dotted lines represent sharp increase between Saturday and Sunday nights. (4) Frequency of heaviest weekly weight on various days of week (5) Frequency of lightest weekly weight on various days of week. (6) Frequency of onset of menstruation on various days of week

cycles of irregular length, variations in food intake, colds, constipation, and other physical and behavioral disturbances were all reflected in weight changes When the same curves were plotted over the days of the week, however, it became evident that the weights of most of the subjects, on most of the weeks of the study, tended to rise sharply at the beginning of each week and then to decrease gradually toward the week-ends The weight curve of Subject 19 (Figure 1, No. 1) is an example of unusually regular weekly fluctuations, though it is not typical of the group because of the slowness and resulting lateness in the week with which this subject frequently attained her maximum weight The upward trend of this girl's weight curve was due to rapid recuperation from a recent illness.

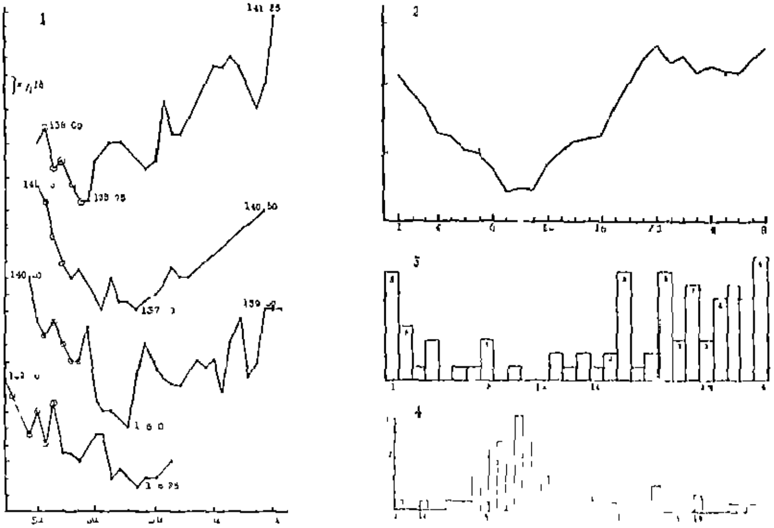


FIGURE 2

Monthly Weight Cycles

(1) Consecutive monthly weight cycles of Subject 1, with highly irregular weekly cycles Circles indicate days of menstruation (2) Composite results secured by averaging deviations from means of 48 cycles, reduced to equivalent of 28 days, and corrected for average loss of ¼ pound per subject per month Curve smoothed by averaging results for each day with those of the two adjacent days (3) Frequency with which heaviest day of month fell on various days of menstrual cycle (4) Frequency of menstrual cycles of various lengths number 85, mean 31 25 days, mode 28 days

Greater, though less regular, weekly fluctuations are shown in the weight curve for Subject 22 (Figure 1, No. 2). A third extreme is shown in the cycles for Subject 1 (Figure 2, No. 1), which were less regular and of more limited range than those of the two subjects just mentioned, though they formed a pattern more typical of the group as a whole in that they tended to rise sharply between Saturday and Sunday nights and to decrease slowly toward weekends. The weekly cycles demonstrated by our other subjects largely fell within the range established by these three cases, with variations in the day of the week upon which the maximum and minimum weekly weights typically fell.

Weekly fluctuations in weight ranged from 0.5 to as much as 6.75 pounds. As shown in Table 1, average weekly changes for the various subjects ranged from 1 5 to 3 29 pounds, with an average for the entire group of 2 11 pounds. In percentage of subjects' average weights the average weekly changes ranged from 1.27 to 2 45 per cent In considering these figures it must be remembered that many of the fluctuations in weight were of sporadic nature, unrelated to the weekly rhythms we are considering. Such fluctuations were as likely to fall on one day of the week as another, and accordingly either to enhance or to reduce the measurable weekly fluctuations

The average trend of the weekly fluctuations for each subject was determined by averaging the weight for all of the Mondays, Tuesdays, Wednesdays, etc., throughout the period of the study. Examples of the curves thus secured are shown as Figure 1, Nos. 3*a, b, c, d*. Of the 28 subjects 6 had the heaviest average weight on Sundays, 4 on Mondays, 10 on Tuesdays, 6 on Wednesdays and 2 on Thursdays. Less variable was the frequency with which the weight of the subjects increased sharply between Saturday and Sunday nights Twenty-two of the 28 subjects exhibited average increases of from 1 to 4 pounds between these nights, 5 increased less than 0 5 pounds, and one had a smaller average weight on Sunday than on Saturday nights.

The data were tabulated further to determine the frequency with which the maximum and minimum weights for all subjects fell on the various days of the week. To do this we tabulated the days of each week upon which each of the subjects weighed the most and the days when they weighed the least. When either maximum or

TABLE 1
Weekly and Monthly Weight Changes of Individual Subjects

Subject No	Age	Average weight	Av. loss at week-ends	Av. loss in % of av wt.	Max weekly loss	Min. weekly Loss	Loss at time of menses	Average monthly loss	Average month-ly loss in % of av wt.
1	17	137.11	1.83	1.33	3.75	.50	2¼, 3¼, 4½, 3¼	3.44	2.51
2	15½	127.47	2.01	1.58	3.25	.75	4½, 2½, 5½	4.08	3.20
3	17	115.32	1.63	1.38	3.25	.75	3¼, 4¼, 1¾, 4¾	3.50	3.03
4	17	116.06	2.23	1.92	3.25	1.25	2¾, 5, 2, 5	3.69	3.18
5	15½	125.13	1.60	1.28	3.00	.75	2½ 2¾	2.62	2.09
6	15½	140.52	2.57	1.83	5.50	1.00	(4½)*, 2		
7	15½	113.41	2.75	2.42	5.00	1.00	(2½, 2½)*		
8	16	102.73	1.78	1.73	3.75	1.00	(3½)*, 2¾, 2½		
9	17	97.84	1.86	1.90	2.75	1.00	2¾, 3	2.87	2.93
10	17½	115.42	1.91	1.65	3.75	.50	2¼, 4¾, 1½, 3	2.87	2.49
11	27	126.16	1.89	1.50	3.25	1.00	2, 1½, 2¼, 2¾	2.10	1.66
12	19	125.41	2.12	1.69	4.75	1.00	2½, ¾, 2½	1.92	1.53
13	17	122.03	2.09	1.71	3.75	.75	(1½, ½)* 5¼		
14	17	128.53	1.70	1.32	3.75	1.00	3, 2½, 6¾, 1½	3.44	2.68
15	17	139.13	2.30	1.65	3.50	1.25	3, 3¾	3.37	2.42
16	28	113.68	1.50	1.32	2.50	.50	2¼, 2¼, 1¼	1.92	1.69
17	14	110.97	1.86	1.68	3.25	.75	3, 3½	3.25	2.93
18	16½	149.91	2.70	1.80	5.25	1.25	5, 4¾, 4	4.58	3.05
19	16½	114.06	1.86	1.63	6.75	.75	(3¾, 1½)*		
20	18	118.20	1.91	1.61	3.50	.75	3, 3¼	3.12	2.64
21	21	119.18	2.21	1.99	4.00	1.00	1¾, 2½ (1¾, 4½)*		
22	18	123.99	2.54	2.05	4.00	1.50	6, 5¾, 6	5.92	4.77
23	16	143.50	1.82	1.27	3.25	1.00	1, 1	1.00	.70
24	17½	118.73	2.50	2.10	4.00	1.00	4½, 3¾, 3½	3.92	3.30
25	18	147.15	1.89	1.28	3.25	1.00	2½, 2¾, 1	2.08	1.41
26	17	130.27	2.46	1.89	4.00	1.50	2½, 2	2.25	1.73
27	17½	134.38	3.29	2.45	6.00	1.75	3¼, 6, 4	4.25	3.29
28	17	155.25	2.17	1.40	5.00	1.00	3, 2½, 6	3.83	2.47

*Figures in parentheses are weight increases at time of menses minimum weight of a subject for a given week was duplicated on more than one day of that week, each of these days was given one count. The results (*Figure* 1, Nos. 4 and 5), indicate the tendency of the maximum weights of the group to fall most frequently on Tuesdays, Mondays, and Sundays, in the order stated, and for the minimum weights to fall most commonly on Saturdays.

As previously stated, our subjects followed similar weekly programs. The average number of hours they spent at school work on the various days of the week was: Monday 8.4, Tuesday 7.3, Wednesday 8.2, Thursday 7.0, Friday 7.7. On these days their activities as well as their retiring and getting-up times were fairly uniform. Saturdays and Sundays, however, were characterized by freedom from scholastic duties, longer periods of sleep, and greater food-intake. The decrease in weights of these young women, which typically occurred during the middle and latter part of the week, apparently reflected the effect of their weekly scholastic program. There then followed the week-end period of relaxation from prescribed routine with increase in food consumption. These factors were reflected in the sharp increases of Sunday nights, from the lows of Saturday nights, in all but one of our 28 subjects. Following this week-end increase there typically were continued but more gradual increases in weight until most of the subjects attained their maximum weights for the week between Monday and Thursday nights, inclusive. This slow continuation of the rise beyond the week-end we believe to be due to increased water retention following the ingestion and assimilation of extra food and the experiencing of extra rest, over the Sabbath period.

If, as we believe, the weekly weight rhythms are the result of the weekly changes in living habits, then similar weekly weight cycles would be expected to occur in men under comparable conditions. According to a report we have received, two male students at Yale University chalked up their daily weights for several weeks on the wall of their dormitory bathroom. The resulting curves exhibited weekly rhythms not unlike those we secured for our female subjects.

In contrast with the weight records of our human subjects were the weight curves secured by Tinklepaugh for chimpanzees. These animals experienced little or no week-end changes in their living habits and likewise exhibited no weekly rhythms in weight.

Available evidence of weekly rhythms in behavioral or physiological functions naturally is slight at the present time, for apparently the existence of such rhythms has seldom been suspected and therefore has not been watched for. We submit examples of the type of evidence we have discovered, hoping that other investigators may carry the search farther by examination of original data of past or future studies.

The term "blue Monday" still has a definite connotation of psychological and physiological depression, even though the burdens of wash day have to a large extent been taken over by our modern laundries A similar phenomenon is frequently found in industrial studies of worker output, where lowered efficiency—"the Monday effect"—is commonly attributed to differences in incitement. Much of the data presented in behavioral studies, when plotted by days of the week, indicates the occurrence of weekly rhythms, though usually the published data are inadequate for such treatment For example, the crude data for the stylus tapping test used by Hollingworth (3) in her study of functional periodicity in women, when plotted, form jagged graphs with saw-tooth patterns not unlike those of our weight curves. When averages for various days of the week are secured, the resulting curves commonly are low on Mondays and tend to rise toward the week-end. The results of the same author's tests of steadiness and color sorting are too variable for such treatment

Evidences of weekly rhythms may occasionally be found in studies of physiological functions. Thus when the crude data of systolic, diastolic and pulse pressures secured by Eagleson (1) for her subject *C*, over 22 weeks, are converted into average weekly curves, each of these functions is high on Sunday and tends to decline toward the week-end. Similar results were not attainable with the data published for her other subjects Sandiford and Wheeler (6) studied the metabolism of one woman through pregnancy and lactation. In the portion of the graph presented by them which covers determinations during the lactation period, during which observations were made three or four times weekly, the metabolic rate is generally higher during the mid-week than on week-ends.

While evidence as limited as the above is not conclusive, it suggests that weekly rhythms in weight are a result of changes in the general functioning of the organism and that the functional changes may be attributed to the weekly living regimen. On the behavioral side we feel justified in assuming that for every physiological change there is a tendency toward a behavioral change, though except in extreme cases, this tendency may be compensated for either by habit or through conscious expenditure of increased or reduced effort by the subject.

We would not expect persons living less routinized lives than our subjects to exhibit as clearly defined weekly cycles. A day off or a night's revelry might easily cause the equivalent of the maximum effect we observed at week-ends. In like manner several such days or nights a month would be expected largely or completely to obscure evidence of any underlying weekly or monthly rhythm

C. MONTHLY WEIGHT CYCLES

Underlying the saw-tooth pattern caused in the weight curve by the weekly cycles, our data indicate that in women there is a monthly weight cycle which is coordinated with the menstrual cycle. In the monthly cycles weight tends to drop sharply through, and commonly beyond, the period of the menstrual flow and then to increase gradually until the high point is again attained at or a short time before the next onset of the flow. In some cycles the peak weight occurs on or about the 21st day (after the onset of menstruation) and decrease continues until Day 5 to 10 of the following cycle, as shown in Figure 1, No. 2 In others the peak is reached at or within a day or two preceding the onset of the flow as shown in Figure 2, No. 1. In a few instances, the only discernible drop occurred during the flow. The frequency with which the maximum weight of the monthly cycle occurred on various days of the cycle is shown as Figure 2, No. 3.

The general trend of the monthly weight cycle is shown in Figure 2, No. 2. This is a composite curve of 48 cycles, 11 of which included one or more unrecorded days. Cycles under 20 or over 36 days in length, or which contained extreme fluctuations apparently not related to the monthly cycles, were not used. Data for the graph were prepared by obtaining the daily deviations from the average monthly weight of each subject to the nearest half-pound. The data were corrected for an average decrease of one-fourth pound per subject per month and all cycles used were reduced to the equivalent of 28 days in length. Averages of the daily fluctuations were then secured for the total number of cycles. The data thus obtained are indicated by dots. The curve was smoothed by averaging the values for each day with those for the one preceding and one following days. Though the curve has a smaller range than those plotted for single cycles, it presents the general trend of the monthly weight cycle. The two peaks reached near Days 21 and 28 of the cycle represent the tendency of the weights of different subjects to reach their maximum on or near one or the other of these two days.

In the monthly weight cycle described by Tinklepaugh for chimpanzee, the maximum weight typically was reached on Day 24 of a 36-day cycle, which is comparable to Day 21 of the 28-day human cycle

The most uniform characteristic in the monthly weight cycles we have described is the tendency of weight to decrease during menstruation. Cases in which menstrual cycles did not occur or in which weight did not decrease during the flow seem significant. Twenty-six of our 28 subjects experienced monthly weight cycles on most of the months they were under observation Cyclic weight changes did not occur in three subjects during amenorrhea. Of 81 menstrual cycles adequately covered by our data, 71 were characterized by weight decreases during and beyond the menses (Table 1). In each of the remaining 10 cycles in which weight atypically increased during the flow, there seemed to be one or the other of two causes responsible During the study a few subjects increased or decreased rapidly in weight during or following minor illnesses. In some cases such trends were not counteracted by monthly fluctuations (Figure 1, No. 1). In a few instances the onset of menstruation atypically occurred at or near the end of the week so that the growth phase of the weekly cycle coincided with the regression phase of the monthly cycle In some such cases the weekly cycle apparently was more effective than the monthly one, resulting in weight increase during the flow (Figure 2, No. 1, Cycle 4).

The gross weight losses during the regression phase of the monthly cycles, measured from the highest point during the premenstruum to the lowest point during the postmenstruum, ranged from 0 75 to 6 pounds Average losses for the various subjects were from 1 to 5 92 pounds, and average losses in percentage of weight ranged from 0.70 to 4.77 per cent The losses in these cycles, as thus measured, apparently were either increased or reduced by the effects of the weekly cycles, depending upon whether regression phases, growth phases, or regression and growth phases of the two cycles chanced to coincide. Menstrual malaise or other physical ailments also affected the cycles at times

Monthly rhythms in various physiological and behavioral functions of women have been observed by numerous investigators, though there have been many inconclusive or conflicting findings. It seems probable, in the light of the present study, that many significant data concerning such rhythms may have been obscured or rendered meaningless because of failure on the part of the investigators to recognize in their results, or to allow in the spacing of their determinations, for weekly fluctuations which may have been as great or even greater than the monthly ones.

For references to original studies we refer the readers to Eagleson (1), who reviewed the literature up to 1924, and to Seward's splendid review, carrying the subject up to 1934 (7). For our purposes here it will suffice to discuss the results in a general way.

Studies of body temperature, blood pressure and metabolism, though frequently inadequate in the number of subjects used, the duration of the studies, or the frequency or timing of determinations, have brought forth much data of a positive nature. In the majority of cases, where results have been positive, the rhythms discovered have followed a pattern of increase before menstruation and of decrease during the menstrual flow. In the interpretation of their data several investigators discounted their results on the grounds that rhythms did not occur uniformly in all subjects. Others combined their data, both positive and negative, and arrived at results that were entirely inconclusive The only general conclusion that can be drawn from the combined results of these investigations is that many women do exhibit monthly rhythms in temperature, blood pressure, and metabolism, that are comparable in pattern to the weight cycles we discovered in our subjects

On the more complex behavioral side as would be expected results are still less conclusive. It is common knowledge that many women experience a three- or four-day period of hyper-irritability just prior to menstruation, while others at that time suffer pre-menstrual malaise with accompanying lethargy. Other women experience similar states during the first days of the flow It is not surprising, then, to find, as Kleinknecht and Schaare (5) did, that some women show increase, others decrease, in errors of perception during the menstrual period, or that Fearing's subject, in a study of static equilibrium (2), became so tense and restless during the premenstrual period as to obscure practice, and that the amount of sway began decreasing at the beginning of flow Johnson (4), using 34 subjects, found diminished efficiency in tight-wire walking during menstruation, with improvement following until about the 11th day after the end of the flow. Hollingworth (3), on the other hand, in her study of a series of tests of motor and mental ability, found no rhythms in efficiency associated with the menstrual cycle It seems possible, however, that tendencies toward lessened effectiveness prior to or during menstrual periods exist, but that subjects compensate for them by putting forth greater effort at that time.

Though we know that there are cyclic changes in both the physical and mental well being of many women, there is little evidence indicating that these cannot be compensated for, so far as behavior is concerned, and such compensation would be most likely to occur under the incentives ordinarily provided by test conditions The fact that the rhythmic pattern of weight in our study is similar to that found in the majority of physiological studies in which the *results have been positive, leads us to believe that the monthly* weight fluctuations are a reflection of these other physiological changes *in the organism. We further assume that there are underlying* tendencies toward increased or decreased behavioral effectiveness *which are correlated with the physiological changes*

D. THE RELATIONSHIP BETWEEN WEEKLY AND MONTHLY CYCLES

The menstrual cycle, when considered in terms of the accompanying changes of the uterine endometrium, consists of a growth and a regression phase Menstruation commonly is considered as terminating the regression phase It is noteworthy that the regression phase of the monthly weight cycle in our subjects begins between Days 21 and 28, or near the time when the regression phase of the uterine cycle is believed to begin. But, as we have indicated, weight loss and, occasionally, decline in some physiological functions, continue beyond the regression phase of the uterine cycle to as late as Day 8 to 12 of the menstrual cycle

Though our evidence of weekly weight cycles was secured from young women living in the thoroughly established weekly routine of a school, we have presented some evidence that similar cycles may occur in the physiological functions of women. When the artificially determined weekly cycles and the spontaneously determined monthly cycles are observed, the weekly cycles seem to be as general in their effect upon the organism as the others. Bearing out this contention, we have shown how, in the case of weight, the weekly and monthly cycles may enhance or neutralize each other when similar or opposing phases of the two cycles chance to coincide.

The interplay of the weekly and monthly cycles, one upon the other, may be responsible for the slight sequential relationship between menstrual onset and peak periods in weekly weight in our subjects. As shown in Figure 1, No. 4, the heaviest day of the week most frequently occurred on Tuesdays, Mondays, and Sundays, in the order stated. The onset of menstruation most commonly fell on Monday, Tuesday, and Wednesday (Figure 1, No. 6). This would seem to indicate that menstrual onset commonly preceded the heaviest day of the week. In reality, however, the heaviest day typically fell a day or so earlier in weeks of menstrual onset than it did in other weeks. The heaviest day of the week preceded onset of the flow in 59 per cent of the cases, coincided with onset in 27 per cent of the cases, and followed onset in only 14 per cent. The explanation of this fact may lie partly in the earlier termination of weekly weight increase because of physical disturbances associated with the onset of the flow On the other hand, the tendency of onset of the flow to occur most often toward the beginning of the week may have been due to acceleration of the growth phase of the menstrual cycle by coincidence of the growth phase of the weekly weight cycle

In the estimation of the authors, the entire subject of weekly and monthly rhythms in both physiological and behavioral functions justifies renewed study under conditions in which both weekly and sporadic daily fluctuations can either be controlled or be recognized and allowed for in the analysis of the data secured. Physically normal subjects, both mature and immature, in institutions where living habits could be varied according to experimental needs, would offer opportunity to settle this important problem conclusively.

E. SUMMARY AND CONCLUSIONS

Twenty-eight young women, including two teachers, living in the relatively uniform environment of a preparatory school, exhibited weekly weight cycles in which there were fluctuations of between 0.5 and 6.75 pounds. Most of the same subjects exhibited monthly weight cycles, correlated with the menstrual cycle, in which weight fluctuated between 0.75 and 6 pounds.

In the weekly cycles, weight typically increased sharply between Saturday and Sunday nights, and reached its maximum one to three days later. Weight then tended to decrease slowly to its weekly low on Saturday night. The abrupt week-end rises are believed to have been caused by increases in rest and food-intake during week-ends.

Limited evidence was found in data of other investigators indicating that unrecognized weekly rhythms at times occurred in blood pressure, pulse, metabolism, and behavioral efficiency.

The monthly weight cycles corresponded closely to the monthly rhythms found by some investigators in the physiological functions of women. Weight typically decreased through periods of menstruation, beginning at or a few days before the onset of the flow, and continuing through or a few days beyond its termination. Weight then rose slowly until the time for the next menstrual onset was approached.

The two cycles, weekly and monthly, tended either to enhance or to neutralize each other, accordingly as their corresponding or opposing phases chanced to coincide.

The authors believe that the artificially induced weekly cycles or the sporadic fluctuations caused by day to day changes in living habits, both of which are reflected in weight changes, may obscure or render inconclusive experimental evidence of spontaneously generated monthly rhythms such as are known to occur in many women. This would be more true in studies of physiological than of behavioral functions, for differences in tendencies toward decreased or increased behavioral effectiveness would be more largely compensated for under the conditions of incentive provided in experimental situations.

The indicated effects of the physiologically determined monthly cycle upon behavior, and of the behaviorally determined weekly cycle first upon physiological functions and then upon behavior, afford complex but enlightening examples of the interplay between physiological and behavioral functions

REFERENCES

1. EAGLESON, HELEN E. Periodic changes in blood pressure, muscular co-ordination and mental efficiency in women *Comp Psychol. Monog*, 1927, **4**, No 20 Pp 65.
2. FEARING, F K The factors influencing static equilibrium *J Comp. Psychol*, 1924, **4**, 91-121
3. HOLLINGWORTH, LETA S Functional periodicity—An experimental study of the mental and motor abilities of women during menstruation Teachers College, Columbia Univ, 1914. Pp viii+101.
4. JOHNSON, G B. The effects of periodicity on learning to walk a tight-wire *J Comp. Psychol*, 1929, **13**, 133-141
5. KLEINKNECHT, F, & SCHAARE, M. Hat die "physiologische Wellenbewegung" einem Einfluss auf die Lageempfindung der Frau? *Zbl. Gynak*, 1028, **523**, 2709-2710
6. SANDIFORD, I., & WHEELER, T. The basal metabolism before, during and after pregnancy *J. Biol Chem*, 1924, **62**, 329-352.
7. SEWARD, GEORGENE H The female sex rhythm. *Psychol Bull*, 1934, **31**, 153-192
8. TINKLEPAUGH, O. L. Sex cycles and other cyclic phenomena in a chimpanzee during adolescence, maturity, and pregnancy. *J Morph.*, 1933, **54**, 521-547

450 West 24th Street
New York City

Psychopathic Hospital
Iowa City, Iowa

The Journal of Genetic Psychology, 1939, 54, 17-25

AGE AND SEX DIFFERENCES IN CHILDREN'S OPINION CONCERNING SEX DIFFERENCES*

Institute of Child Development, University of Washington

STEVENSON SMITH

Among the most compelling of objective conditions under which human beings live is the opinion of others concerning them. If it is generally believed that men drive cars better than women do, or that men are by nature better fitted to become surgeons or business executives, then these beliefs make women's vocational achievement more difficult, regardless of whether the beliefs are true or false. Because opinion is always an objective social factor, the measurement of opinion has an importance that is quite apart from the importance of the opinion's being true.

In the study that is here reported 100 boys and 100 girls in each year-of-age group from 8.0 to 15.0 years voted as to whether boys or girls possess to a greater degree each of 19 desirable traits and 14 undesirable traits. The subjects were all pupils in the coeducational grade schools of Seattle. Whether each trait would be considered desirable or undesirable by children was determined by the independent vote of 17 experienced teachers. Their opinion was in complete agreement.

The questions were read to classroom groups in the following form: "*Who are* . . . (e.g., kinder to animals) . . . , *boys or girls?*" Each child was provided with an answer blank on which were printed the numbers of the questions and following these the letters *B* and *G* (e.g., 1 . . . *B* . . . *G*). After general directions had been read, preliminary practice was given on obvious material, such as "*Who usually have longer hair, boys or girls?*" The children responded by circling the *B* or the *G*.

Table 1 gives the list of traits, desirable (*D*) and undesirable (*U*), in the order in which the items were presented, together with the boys' and girls' votes at the various ages. These vote entries are the proportion of voters who said that *boys* rather than girls possessed

---

*Received in the Editorial Office on March 14, 1938.

TABLE 1

VOTE OF BOYS AND OF GIRLS THAT BOYS PREDOMINANTLY POSSESS THE TRAITS LISTED

	Trait		Boys' vote at these ages							Girls' vote at these ages							Boys	Girls
			8	9	10	11	12	13	14	8	9	10	11	12	13	14	8-14	8-14
1	Cry when hurt	*U*	08	02	03	00	00	01	04	01	02	04	04	02	02	00	026	021
2	Bossy	*U*	72	64	53	59	56	50	42	72	73	69	54	52	43	45	566	583
3	Kind to animals	*D*	31	23	23	40	42	47	36	18	11	10	18	15	15	13	346	143
4	Do things for others	*D*	50	61	58	61	63	68	66	27	27	23	29	29	23	23	610	259
5	Do right unobserved	*D*	40	50	33	42	26	32	33	15	08	11	16	10	08	11	366	113
6	Quick in movement	*D*	90	91	90	95	94	94	94	83	88	84	84	81	77	77	926	820
7	Tire easily	*U*	13	11	07	03	01	01	04	06	11	07	02	01	05	01	057	047
8	Quick to anger	*U*	78	72	64	56	49	58	40	81	76	66	51	54	47	40	596	593
9	Afraid of strangers	*U*	03	03	04	02	01	02	07	03	03	02	05	03	08	08	031	046
10	Memory for stories	*D*	61	59	62	63	63	55	64	36	31	35	40	39	48	48	610	396
11	Mean to strange dogs	*U*	72	82	80	73	68	70	73	92	95	90	87	87	85	89	740	893
12	Hard workers	*D*	80	73	87	97	90	89	83	42	62	63	57	65	59	68	856	594
13	Quarrelsome	*U*	62	63	67	52	40	46	36	80	74	70	52	58	45	40	523	599
14	Easily frightened	*U*	04	01	01	00	02	01	00	06	08	02	00	02	02	00	013	029
15	Selfish	*U*	64	58	51	37	38	39	37	93	86	79	68	57	55	47	463	693
16	Like to tease	*U*	79	82	81	78	86	86	76	97	94	98	99	95	96	97	811	966
17	Friendly	*D*	23	31	36	49	53	54	71	01	08	07	10	19	17	27	453	127
18	Mean remarks	*U*	79	74	78	68	74	52	43	94	95	96	87	65	52	45	669	763
19	Endure pain	*D*	94	97	1 00	97	1 00	97	92	88	90	93	98	93	86	85	967	904
20	Destroy property	*U*	89	96	90	94	96	95	95	98	98	98	99	99	1 00	99	936	987
21	Stick to hard job	*D*	88	90	92	96	87	87	90	72	61	68	69	81	68	65	900	691
22	Keep secrets	*D*	62	56	56	67	80	75	77	11	16	32	27	61	60	61	676	383
23	Best leaders	*D*	73	74	84	91	89	89	96	47	45	62	75	74	72	81	851	651
24	Honest	*D*	29	30	34	31	30	23	38	07	08	10	07	07	08	18	307	093
25	Nice to small children	*D*	21	14	21	16	12	08	12	01	03	03	01	03	02	04	149	024
26	Likely to cheat	*U*	74	72	80	72	83	82	79	92	92	90	90	78	78	75	774	850
27	Good losers	*D*	46	51	49	60	67	62	71	37	38	33	40	58	55	53	580	449
28	Truthful	*D*	26	35	32	38	30	24	25	05	10	08	07	12	10	17	300	099
29	Puzzle solvers	*D*	81	81	83	82	81	62	80	62	74	70	63	63	51	61	786	634
30	Generous	*D*	41	32	34	38	47	46	54	09	10	12	11	16	23	35	417	166
31	Afraid of the dark	*U*	04	03	02	01	00	02	04	06	06	08	02	05	04	02	023	047
32	Good sports	*D*	75	80	76	81	83	79	77	41	46	46	61	71	55	58	787	540
33	Bright in school	*D*	37	49	46	39	32	34	25	05	09	10	09	13	18	17	374	116
	All 19 *D* traits		552	567	577	623	615	592	623	319	339	358	380	426	397	433	593	379
	All 14 *U* traits		501	488	472	425	424	418	386	586	581	556	500	470	444	420	444	508

TABLE 2

PROPORTION OF VOTE FAVORING THE GROUP JUDGED

(19 good traits, 14 bad traits.)

	8	9	10	11	12	13	14	All ages 8 14
Good traits								
B opinion of *B*	.552	.567	577	623	.615	592	623	593
B opinion of *G*	448	433	423	377	.385	408	377	407
G opinion of *B*	319	339	.358	.380	.426	397	433	.379
G opinion of *G*	.681	661	.642	620	574	603	567	621
Bad traits								
B opinion of *B*	499	512	.528	575	576	582	614	556
B opinion of *G*	.501	488	.472	425	424	418	386	.444
G opinion of *B*	414	419	.444	500	530	556	580	492
G opinion of *G*	.586	581	556	500	470	444	420	508
Good and bad traits								
B opinion of *B*	529	544	556	602	598	.588	619	577
B opinion of *G*	471	456	444	.398	402	412	.381	423
G opinion of *B*	361	373	394	431	470	465	495	427
G opinion of *G*	639	627	606	569	530	535	505	573
Good and bad traits								
B and *G* opinion of *B*	445	458	.475	516	534	526	.557	501
B and *G* opinion of *G*	555	542	.525	484	466	474	443	499

each trait. The proportion of voters who assigned each trait to the girls is, therefore, 1 minus the entry. Likewise, in converting the measure of an undesirable trait into one that can be pooled with the measures of desirable traits we subtract the entry from 1. If, for example, 60 per cent of the girls in any age group were to vote that it is the boys who predominantly possess a certain undesirable trait, then 40 per cent of these girls' vote would be "favorable" to the boys.

The results here reported are not measures of the degree of conviction of the individuals who make up each voting group. They are merely statements of the proportional vote of the groups. Also this is not an investigation of opinion concerning faults and virtues in general but only of opinion concerning sex differences in traits as listed in the study. When we use such a statement as "boys regard themselves as superior to girls in good traits" we mean to imply the more precise statement, "the summation of the boys' votes as to which sex possesses the desirable traits above listed shows a majority for boys." This reservation will be taken for granted in what follows so as to facilitate description

The main tendencies as expressed in Table 1 may be summarized as follows.

1. With increase in age the boys have a progressively poorer relative opinion of the girls and the girls have a progressively better relative opinion of the boys (Figure 1). Conversely, with increase in age the boys have a progressively better opinion of themselves and the girls have a progressively poorer opinion of themselves Both girls and boys, that is, have a progressively better opinion of the boys

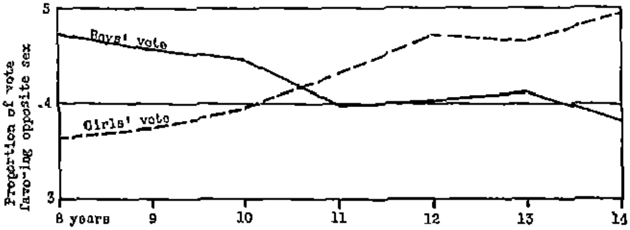


FIGURE 1

BOYS' AND GIRLS' VOTE THAT THE *opposite sex* IS MORE POSSESSED OF THE 19 DESIRABLE TRAITS AND LESS POSSESSED OF THE 14 UNDESIRABLE TRAITS, *DT* AND *UT* COMBINED

Each year of age group is indicated by the lower limit of the year interval.

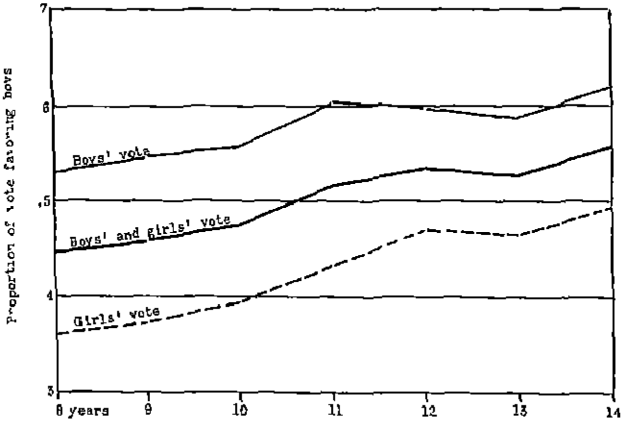


FIGURE 2

BOYS' AND GIRLS' VOTE THAT BOYS ARE MORE POSSESSED OF THE 19 DESIRABLE TRAITS AND LESS POSSESSED OF THE 14 UNDESIRABLE TRAITS, *DT* AND *UT* COMBINED and a progressively poorer opinion of the girls (Figure 2). The vote proportion is determined by the purely accidental selection of the traits on which the voting was done. If another array of traits were selected at random any vote proportion progressively favorable to boys would be different and, though far less likely, the proportion *might progressively favor girls. The probability that the* latter might happen is negligible provided the selection of traits was, as intended, a purely chance selection.

2. Although each sex thinks better of itself than it does of the other sex, at fourteen years the girls have reached a point where they think almost as well of the boys as they think of themselves, as far as this array of desirable and undesirable traits is concerned. Both boys and girls are more likely to assign good traits to their own sex than to assign bad traits to the other sex (Table 2 and Figure 3) They show, that is, more humility in judging themselves on the basis *of bad traits and less humility in judging themselves on the basis of* good traits. This tendency, generally greater in the girls, increases with age in the girls and decreases with age in the boys By the age of eleven and a half the majority of girls begin to see more of the listed faults in themselves than they do in boys, but even up to fourteen and a half years they see more of the listed virtues in themselves than in boys. At fourteen and a half the boys regard *themselves as about equally possessed of these virtues and free from* these faults. In the vote of both boys and girls as age increases there is a greater change in the sex allocation of the undesirable traits than of the desirable traits (Figure 3) The two scatter diagrams (Figure 4) show that the proportion of boys' votes agrees better with the proportion of girls' votes in the array of undesirable traits than in the array of desirable traits. Also the vote deviates more from the .50-.50 proportion *in the undesirable than in the desirable traits.* Corresponding to this, the correlation between boys' and girls' opinion is higher in the allocation of undesirable traits.

All these measures are relative and do not imply the absence of an absolute improvement in both sexes as age increases The measures do imply that, if there is absolute improvement in these traits, the improvement is greater in boys according to the opinion of both boys and girls. This does not mean, of course, that their opinion is correct.

The increasing favorableness toward boys in the girls' vote as the

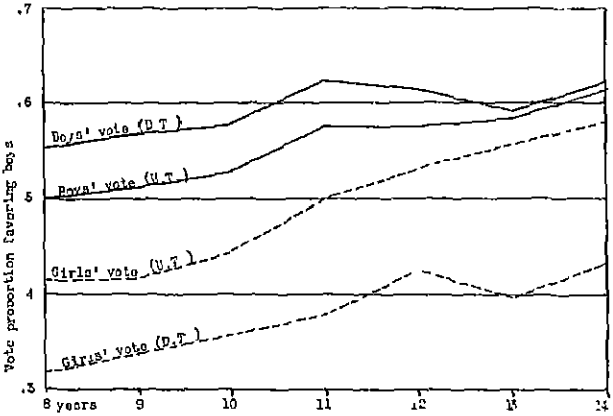


FIGURE 3
Boys' and Girls' Vote That Boys are More Possessed of the 19 Desirable Traits and Less Possessed of the 14 Undesirable Traits, *DT* and *UT* Considered Separately

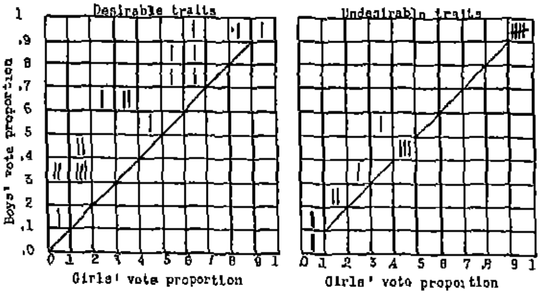


FIGURE 4
Correspondence of Boys' Vote to Girls' Vote. Vote Favorable to Boys, All Ages Combined

Scatter diagrams showing that boys and girls agree more closely in allocating the undesirable than in allocating the desirable traits to either sex.

girls grow older may correspond to a general improvement in social comprehension that makes the girls more and more free from prejudice. Possibly boys by the age of eight years already have a less prejudiced understanding of sex differences than girls have, and possibly the boys' increasingly poor opinion of girls is evidence of the growth of those reactions which constitute what is called "masculine protest" But such speculations have little value in the absence of any measures of the concomitant factors that such an hypothesis assumes.

The pooled opinion of both sexes and all ages is overwhelmingly that girls, rather than boys, are bright in school, cry when hurt and cannot endure pain, are easily frightened, especially by the dark and by strangers, are kind to animals and to small children, are not given to teasing or destructiveness, tire easily, move slowly and do not stick to a hard job, are truthful, honest, not likely to cheat, and will do the right thing even though not watched. The boys in comparison to the girls show the reverse of this picture. That this is not a complete picture of sex differences in personality is obvious enough. It was not the purpose of this study to analyze the personality makeup of boys and girls, but rather to sample a few traits in order

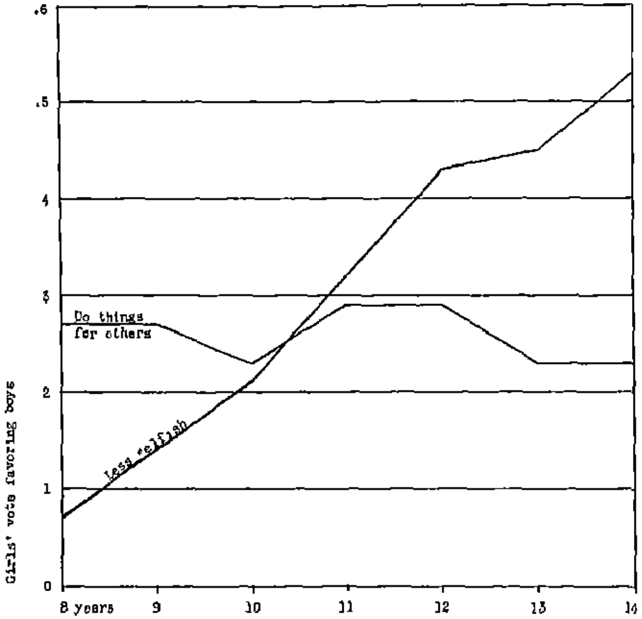


FIGURE 5

GIRLS' VOTE THAT BOYS ARE LESS SELFISH (ITEM 15) AND THAT BOYS ARE MORE LIKELY TO DO THINGS FOR OTHERS (ITEM 4) to obtain some knowledge of the shift of opinion in the two sexes with increase in age.

It is interesting that there is but little relation between certain items which are designated by terms that might suggest a closer correspondence. For example, Item 4, "Doing things for others" has little relation to Item 15, "Selfishness" (Figure 5). Girls' opinion throughout the age range from 8 to 14 concerning boys' selfishness shows a remarkably rapid change which is consistently progressive, so that although 93 per cent of the eight-year-old girls say that boys are more selfish than themselves, only 47 per cent of the fourteen-year-old girls say that boys are more selfish. Throughout this age range there is no consistent change of opinion on the part of the girls concerning boys' tendency to do things for others. Obviously the absence of selfishness and the presence of doing things for others mean very different things.

The reliability of each vote proportion is satisfactorily high. The *PE* of any entry is $.67\sqrt{\frac{pq}{n}}$—and this never exceeds .033 in the traits taken separately and never exceeds .006 when the vote on all the traits is pooled. The writer realizes that such a method of computing reliability is strictly valid only when the voters are wholly unselected. The opinion of a classroom group is subject to the effect of any chance incident, such as a girl's cheating or a boy's kicking a dog, which is known to the group as a whole. As each age-sex group was drawn, however, from at least five classrooms and from at least three schools any constant error of this sort must have been small.

In no trait do the girls, as age increases, receive from the two sexes combined a progressively favorable vote proportion that is statistically significant. On the other hand, if we compute the significance of the progressively favorable combined sex vote for boys in the 33 traits combined (see Figure 2) by comparing the proportional vote favoring boys in the age range 8.0 to 11.0 years (.459) with that in the age range 12.0 to 15.0 years (.539), we find a difference of .08 which is 235 times as large as .00034 which is the *PE* of the difference, where *N* in each group is 19,800 votes. If, in order to avoid halo effect within the individual voter we reduce *N* to 600, that is, to the number of voters, *d* is still 4.1 times the *PE* of the difference. The difference between the eight-year-old and the fourteen-year-old vote proportions favoring boys is .112 and the *PE* of the difference is .024 if N is the number of voters, or .009 if N is the number of votes.

If the traits listed constitute a fairly good chance sampling, that is, if they do not fall into too large clusters and if there is no constant factor operating that caused traits to be selected in which boys appear to show greater improvement than girls, there would seem to be a good probability that in this age range the proportion of both sexes, whose opinion favors boys, varies directly with age

Institute of Child Development
University of Washington
Seattle, Washington

The Journal of Genetic Psychology, 1939, **54**, 27-71

THE ATTITUDES AND INTERESTS OF PREMENARCHEAL AND POST MENARCHEAL GIRLS*[1]

Department of Psychology, Stanford University

CALVIN P. STONE AND ROGER G BARKER

In infra-human primates, as in lower orders of vertebrates, there are both external and intrinsic factors which condition and regulate sexual display, courtship, mounting, and other distinctive patterns of sexual behavior, without which the procreative functions could not be accomplished (4)

In some species of animals characteristic patterns of behavior appear *de novo* when puberal changes in the primary and secondary sexual characters and accessory organs of sex are most in evidence, in others, notably the primates, a limited amount of sexual play is said to appear before the puberal stage of development has been reached In either case, however, observations permit us to assume that sexual drive is greatly augmented near the time of somatic puberty and in most, if not all, instances continues to wax in strength for some time thereafter by virtue of factors of maturation and sexual experience Man, it is popularly believed, affords no exception to these basic generalizations; however, just as closely related species differ from each other with respect to specific details, so also man, with his complicated cultural heritage and social environment, may differ greatly even from the higher anthropoids which most nearly resemble him in structural and physiological organization.

Although this popular belief is one of long standing, scientists, so far, have made little progress in identifying new patterns of behavior in man that manifest themselves at puberty or new urges that are markedly augmented at this time. Moreover, little has been

---

*Received in the Editorial Office on March 20, 1938

[1] Financed by a grant from the Committee for Research on Problems of Sex, National Research Council The junior author, Dr. Barker, is primarily responsible for the statistical analysis of the data of this report We are greatly indebted to Dr Virgil E Dixon, Supt of Schools at Berkeley, California, for permission to use pupils of two of the Junior High Schools of that city and for his generous assistance in compiling data from intelligence tests previously administered in these schools

accomplished in developing methods of investigation that are likely to prove fruitful for the investigation of these or closely related problems Investigators are still exploring for promising leads It is with this type of exploratory research that the present study is to be identified Our study has as its chief objective an inquiry as to whether premenarcheal girls differ significantly from post menarcheal girls of the same chronological ages and social status in their responses to questions concerning their attitudes and interests.

Onset of the menses denotes neither the beginning nor the end of pubescent development. As a rule, prior to the menarche the feminine contour is becoming apparent, there is accelerated increase in height, weight, chest width, and pelvic diameters, the mammary glands enlarge and their contours change, hirsuites in the pubic and axillary region may appear; and the genitalis begin to take on the characteristic adult-like proportions (8). After the time of the first menarche all of the foregoing continue their development for various periods of time In view of these facts, when one compares premenarcheal girls with post menarcheal girls one merely contrasts girls who, on the average, are relatively less mature in physiological sexual development with those who are relatively more mature For our purposes, no serious error is committed if one regards the menarche as a sort of mid-point in the total puberal phase of adolescent physical development.

A THE SUBJECTS

1. *Source, Nationality*

The subjects were 1,000 girls of two large junior high schools of Berkeley, California All of the girls of these two schools were subjected to the same routine of testing and physical measurements Information, supplemented by records from the schools, was secured as to the race of each of the pupils All records from Hebrews, Orientals, South Europeans, Negroes, and Mexicans were discarded. The elimination of these racial stocks appeared to be desirable as it is commonly believed that American-born children of such racial groups differ in age at the menarche and in physical traits from those of middle and north European stock. Moreover, in these school districts such minority groups probably differ greatly in cultural background from the groups who constitute the bulk of our subjects.

Those whose records remain in the study are American-born children of middle and north European stock, and with but few exceptions they have lived in California all their lives.

2. *Chronological Ages*

The number of subjects at each age for whom complete data were secured are shown in Table 1.

TABLE 1

NUMBER OF SUBJECTS FOR WHOM COMPLETE DATA WERE SECURED AT EACH CHRONOLOGICAL AGE

	Ages in months										
	133-*138*	139-*144*	145-*150*	151-*156*	157-*162*	163-*168*	169-*174*	175-*180*	181-*186*	Total	Mean age
Post menarcheal	1	8	28	54	90	94	124	119	46	564	167.4
Premenarcheal	11	31	106	87	63	49	24	13	3	387	155.0

As stated above, we have dealt only with junior high school populations, i.e., the seventh, eighth, and ninth grades. Practical considerations made it unfeasible to bring into a study of this kind all of the school children of the age range in which the premenarcheal-post menarcheal differentiation occurs. Factors associated with school achievement leave some of the mentally retarded children in the upper grades of the elementary schools, while some of the more accelerated children have been pushed ahead to the senior high school. Thus at both extremes of intelligence some selection has taken place; by virtue of this we cannot expect to have complete distributions with respect to variability in any of the functions considered in the study.

The great difference in the mean chronological ages of the post menarcheal and the premenarcheal groups made it necessary to distribute the individuals into subgroups with relatively narrow age ranges, since some of the functions to be considered hereafter are closely related to chronological age. Even when making comparisons based on the narrow age ranges given in Table 1 it was important to know whether the mean ages of the post menarcheal and the premenarcheal subgroups were approximately the same. Information on this point will be provided in connection with a subsequent section that deals with our final age groups (cf. p. 39).

3. *Establishing Age at Menarche*

A woman assistant ascertained from each individual whether or not the menarche had been reached and, if so, the age at the time of the first mensis. This inquiry was made at the time the assistant was making the physical measurements [the results of which have been reported elsewhere (10)] and was directed to the student in a matter of fact way in terms which were commonly used in the physical education and hygiene courses given in these schools In only rare instances were explanatory terms required to make the student comprehend what information was desired The examiner asked for the year and the month when first menstruation occurred The student was urged to fix this time in relation to some easily established landmark such as school class, season of the year, birthday, etc , as an aid to establishing the date of first mensis with a minimum of error. In a few instances, when the child was uncertain or frankly did not remember the desired information, she was asked to seek aid from her mother and bring the information on the following day, this, with a few exceptions, was done

We are of the opinion that this information on menarcheal age is not subject to the errors sometimes arising when the children are asked to make all of the necessary computations involved in fixing the date of the menarche In the present case the girls were asked to utilize cross checks to establish the time, and the assistant aided or checked the child in her calculations. Also, in evaluating the accuracy of these data it is important to know that in both of these junior high schools there were well organized departments of physical education in which the didactic phases of the course dealt with various aspects of feminine hygiene No doubt it was because of this latter fact, in part, that the girls were not embarrassed by the questions asked by our assistant, were not reticent about giving information, and, with only a few exceptions, appeared to have the desired information well established in mind

Since in this study our plan is to compare the attitudes and interests of premenarcheal and post menarcheal girls of the same chronological ages, it is inevitable that comparisons at each chronological age will be made between groups with different menarcheal ages. The post menarcheal girls of each age group will have, on the average, a menarcheal age that is lower than that which the premenarcheal group will have, and there will be but little overlapping in the distributions. At the two extremes the differences in menarcheal ages will be greatest. On the basis of the data of Pryor (9) and that of the present study, it has been estimated that a difference of approximately two years in menarcheal age will obtain for our extreme age groups and a difference of from one to 1 5 for the middle groups The exact amounts cannot be determined, of course, without information as to the actual age at which the nonpubescent girls reach the menarche

4 *Socio-economic Status*

Ideally one would attempt to secure data for a study of this kind from pubescent and nonpubescent girls who, with respect to socio-economic status, typify a well defined, homogeneous population. Practical difficulties associated with securing such a group without at the same time losing other desirable features were such that we were forced to choose our groups largely from the standpoint of securing adequate cooperation for our test program from the school authorities. It so happens that our groups were urban residents of a non-industrial, university community wherein a fairly high degree of commonality of the cultural milieu may be said to exist

In order that the socio-economic status of the subjects might be roughly described and that possible differences between premenarcheal and post menarcheal groups might be detected, items relating to economic status were included in our questionnaire (see cover page of the test, page 63). No field check on the accuracy of this information was made It is presumed, however, that the data conform in this respect to other data similarly obtained in connection with educational and psychological tests, since the mode of administration was similar to that of standard group tests with which the children were very familiar.

In analysing these data we placed the subjects in three socio-economic categories (high, intermediate, and low) on the basis of the occupations of the supporting member of the family (usually father), and upon the basis of certain aspects of the home situation The occupations were grouped as follows: (*a*) professional, upper class and middle class occupations, (*b*) skilled occupations, (*c*) semi-skilled and unskilled occupations These groups correspond to the categories of the Goodenough occupational ratings (2) in the following ways (*a*) corresponds to Categories 1 to 4; (*b*) corresponds to Category 5; and (*c*) corresponds to Categories 6 and 7. A few occupations reported by our subjects were not included in the Goodenough scale These were placed in categories which, according to our judgment, conformed to the ideas embodied in the Goodenough scale. The home situation data are taken from the last five items at the top of the cover page of the attitude-interest test (cf page 63). Home situation ratings were given in accordance with whether subjects indicated the presence of the following in the home: (1) a telephone, (2) a daily newspaper, (3) magazines taken regularly, and (4) one or less persons per room living at home

In the high socio-economic category were placed those subjects falling in occupation group (*a*) who also indicated the presence of three or four of the above home situation items In the intermediate category were placed (1) those subjects falling in occupational category (*a*) who indicated the presence of two or fewer of the home situation items, and (2) those subjects of occupational group (*b*) who indicated the presence of more than two of the home situation items. In the low socio-economic category were placed (1) all subjects of occupational group (*c*) and (2) those subjects of occupational group (*b*) who indicated the presence of two or fewer of the home situation items The frequency distributions for the three groups are given in Table 2.

TABLE 2
DISTRIBUTION OF SOCIAL-ECONOMIC CLASSES

	Post Menarcheal	Premenarcheal	Total
High	373	250	623
Intermediate	185	119	304
Low	23	18	41
No data	1		1
Total	582	387	969

The socio-economic class frequencies in premenarcheal and post menarcheal subgroups closely approximate those expected upon the basis *of the distribution in the total group. The* computed chi square indicates that the obtained deviations from the expected frequencies would arise by chance in from 80 to 90 cases in 100 In view of this fact we may proceed in our study with the assumption that little or no difference exists between the premenarcheal and post menarcheal groups with respect to socio-economic status, and that this factor need not be taken into account in our further comparisons of these groups.

It is important, however, to keep in mind the fact that our total group is not representative of the socio-economic *distribution of the* general population. Over 95 per cent fall into occupational categories that are above the level of the semi- and unskilled classes, and all are from urban families. Because of this fact one must refrain from any generalizations concerning the relationship between menarcheal status and socio-economic status in general.

5. *Grade Placement*

The grade placement of the subjects, classified by chronological age and menarcheal status, is given in Table 3. The pubescent girls are accelerated in comparison with the nonpubescent girls. The mean *acceleration varies from .07 to .50 of a grade and the critical ratios* of the differences vary from .83 to 3.07, for the five age groups with sufficient numbers for a statistical test. The probability that five such differences might successively arise from chance is less than 1 in 1000 (5, p. 337). Owing to the selection of our subjects, however, one cannot generalize from these data concerning the relationship between menarcheal status and grade placement of junior high school girls.

6. *Intelligence Quotients*

The data concerning intelligence were secured from the records on file in the Berkeley schools. The tests from which the *IQ*'s were determined and the number of cases determined by each are as follows

Terman Group	511
Otis Self-Administering	8
Kuhlmann-Anderson	435
Stanford-Binet (1st revision)	12
	966

Table 4 gives the distributions of *IQ*'s in the premenarcheal and post menarcheal groups. It will be noted that the mean *IQ* of the pubescent girls is higher than that of the nonpubescent girls at the

TABLE 3

GRADE PLACEMENT FOR AGE AND MENARCHEAL STATUS SUBGROUPS

	Chronological ages																	
	131-138		139-144		145-150		151-156		157-162		163-168		169-174		175-180		181-186	
Grade	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre
L 7	1	11	4	27	15	63	15	21	7	8	1	4	3	4			1	
H 7			1	3	8	30	10	41	9	6	5	5			2		1	
L 8			2	1	4	12	19	20	45	34	27	14	11	5	3	2	1	1
H 8			1		1	1	9	5	15	12	28	15	23	5	8	3	3	1
L 9							1		13	3	30	11	58	8	56	5	16	1
H 9									1		2		29	2	49	3	24	
Number	1	11	8	31	28	106	54	87	90	63	93	49	124	24	118	13	46	3
Mean Post	7.25		7.75		7.59		7.98		8.37		8.72		9.14		9.37		9.38	
Mean Pre	7.25		7.33		7.52		7.80		8.22		8.49		8.64		9.10		8.75	
Diff $M_{Post}-M_{Pre}$					0.07		0.18		0.15		0.23		0.50					
σ_{diff}					0.085		0.085		0.080		0.100		0.160					
C.R.					0.83		2.12		1.85		2.30		3.07					

five ages where there are adequate numbers for satisfactory tests by amounts ranging from .5 to 5.0 points. The differences are from .28 to 2.34 times their standard errors. When the probabilities of such differences are summated over the five categories (5, p. 337), there is a probability of .012 that the five differences might arise by chance. This difference is of the same order of magnitude and statistical significance as that previously reported by us (11).

From a consideration of Table 4, it is clear that there has been selection for relatively high intelligence at the young ages, for relatively low intelligence at the old ages, and for intermediate ability at the central ages. This, as was mentioned on page 29, is the inevitable consequence of the method of choosing our groups; i.e., taking all of the girls in two schools, each of which included but three grades. Because our primary purpose was to study attitudes and interests, many of which are closely related to the social activities of the individuals, we considered it essential to secure subjects from as socially homogeneous a group as possible. This could be accomplished only at the expense of such selection as is pointed out above, for to have secured an unselected sample of girls of the ages in whom the premenarcheal-post menarcheal dichotomy appeared we should have had to secure additional subjects from both the elementary grades and the senior high schools. This was undesirable because their respective social milieux differ so greatly from that of the junior high school. It seemed that our purposes would be served better by limiting the study upon the side of comprehensiveness in order to secure an advantage on the side of control of social activities of the groups to which the subjects belonged.

7. *Physical Status*

A study of the physical development of these premenarcheal and post menarcheal girls shows that when those of the same chronological ages are compared, changes in the relationships between measures of physique and menarcheal age are found, and that the nature of these changes varies with chronological age (10). Our own previous studies show that at all ages the post menarcheal girls are taller, on the average, than the premenarcheal girls, but that the superiority in height gradually diminishes with increasing age and finally, as the terminal stage of growth in height is reached, the late maturing girls become the taller (1, 10). In the case of weight, chest width, and bi-iliac diameter, the means are significantly higher for post menarcheal than for premenarcheal girls. These differences diminish in amounts but persist even at the terminal stages of feminine growth.

B. INTEREST-ATTITUDE TEST

1. *The Test*

The test as administered is reproduced in reduced form in the appendix of this report. In assembling the items for this questionnaire we were guided materially by two considerations:

(*a*) Preliminary work with the Pressey (6) and the Sullivan (12) tests had revealed a considerable group of items which differentiated premenarcheal and post menarcheal subjects of the same chronological ages fairly satisfactorily (11). These items concerned play interests, vocational interests, reading interests, fears, worries, and ideas of right and wrong Pressey and Sullivan had found that chronological age groups could be differentiated on the basis of responses to these same items. Using the items suggested by our first study as a point of departure, we added items in those regions of interest which appeared most promising These items were secured by searching through all of the questionnaire-type personality tests we could locate (in 1934) which had proven suitable for the ages we were to use

(*b*) In addition to following suggestions gained from the preliminary experiment (11), we assembled items bearing upon certain hypothetical premenarcheal-post menarcheal differences which appeared promising, e g, heterosexual interests, family adjustments, interest in adornment, play activities, vocational choice, etc These groupings of items will be considered later in discussing the results To secure such items other tests were scrutinized, school counsellors were interviewed, and groups of girls in a neighbouring high school were asked by their counsellors for suggestions[2]

2 *Test Administration*

The children were examined in groups An experienced test administrator (woman) conducted the test program with the aid of such assistants as she required from our testing personnel and from the staff of the school.

---

[2] For advice and aid in this task we are especially grateful to Miss Virginia Block, then a Research Assistant in the Stanford School of Education, and Mrs Louise Barker of the Redwood City High School

The children were informed that there was no time limit for the test; they were instructed to work until all items had been checked The seating arrangement was such as to minimize the temptation to consult with each other or to answer items as others were answering them. The test administrator emphasized particularly the point that answers were to be tabulated only by people who did not know the students personally and that teachers and parents would not have access to the individual blanks The girls were urged to consult with the assistants, who were stationed in the room, if they were in doubt as to what they were to do or were uncertain as to the meaning of any question. Standardized answers or instructions respecting such matters were given by the assistants

On the whole the pupils appeared to cooperate in the desired manner A few individuals were amused by certain items, and also a few were unwilling to check all of the items because they regarded their own interests and attitudes as highly personal and not subject to interrogation by anyone. How many silently rebelled but checked the items anyway we have no way of telling

3 *Analysis of Test Data*

All data were placed upon Hollerith cards which were classified by the age and pubescence status of the subjects. The number of cases in each group for whom complete test data were secured is shown in Table 1 It was apparent from the beginning that the number of cases in the extreme age groups was insufficient to give reliable differentiations between premenarcheal and post menarcheal girls with respect to frequency of responses For this reason the data of age group 133-138 months were discarded, and those for age groups 139-144 and 145-150 were combined to form Group I The data for age groups 169-174 and 175-180 were also thrown together to form Group V and the three premenarcheal subjects of age group 181-186 were placed in this latter group Groups II, III, and IV are, respectively, the same as the remaining age groups given in Table 1.

The mean ages of the premenarcheal and post menarcheal subjects

TABLE 5

MEAN AGES IN MONTHS OF POST MENARCHEAL AND PREMENARCHEAL GIRLS IN THE FIVE AGE CATEGORIES USED IN ITEM ANALYSIS

Age category	I	II	III	IV	V
Post menarcheal	147 33	154 20	160 20	166 04	174 85
Premenarcheal	146 99	154 05	160 05	165 62	174 82

in the five age groups are shown in Table 5. These means were computed from the original data giving the age to the nearest month At all ages the post menarcheal girls are slightly the older, but the amounts of the differences, from 5 to 15 days, would appear to be quite insignificant for the functions involved in this study.

The frequency of occurrence of each response[3] *to each item* was determined for each subgroup separately and converted into a percentage of the total of such possible responses The standard errors of the differences in the percentages of premenarcheal and post menarcheal girls making each response at each age were then computed for each of the five age groups separately and the probabilities that the differences in the per cents were true differences were determined [Kelly (3), Formulae 40 and 140]. For certain purposes it was obviously desirable to reduce the five values to a single value This we have done by a method proposed by Pearson (5) The assumption underlying the application of this method in the present case is that if the proportions of the premenarcheal and post menarcheal groups making a particular response were, in fact, random samples from the same population, then the difference in the proportions divided by the standard error of the difference (critical ratio) should be a random sample from the "Z" distributions of "*Student*" The procedure for determining the probability that a group of five such differences might arise by chance on the above hypothesis involves the determination of the continuous product of the individual probabilities and reference to a table of probabilities provided by Pearson (5, p 337) For the present purpose the normal curve was used as sufficiently representative of the "Z" curve. In the later presentation of these results, the probabilities have been converted into standard deviation units

---

[3] We have not carried the questioned (uncertain) responses through this step of the analysis, as the frequency of such responses was so small in most cases that the differences were unreliable

As a check upon Pearson's method of combining the data from the separate age groups, and as a check upon the computations, we have drawn the data cards of 201 premenarcheal and 201 post menarcheal subjects, matched with respect to chronological ages but otherwise randomly selected, and determined the significance of the differences in the proportions of the two groups making each response. The distribution of the numbers of cases at each chronological age is shown in Table 6.

TABLE 6
FREQUENCY OF SUBJECTS AT EACH AGE FOR THE MATCHED GROUPS OF 201 PREMENARCHEAL AND 201 POST MENARCHEAL GIRLS

	Chronological ages								
	133- 138	*139-* 144	*145-* 150	*151-* 156	*157-* 162	*163-* 168	*169-* 174	*175-* 180	*181-* 186
Post menarcheal	1	3	23	44	55	37	22	13	3
Pre-menarcheal	1	3	23	44	55	37	22	13	3

In the case of certain items, both premenarcheal and post menarcheal groups favored a response so overwhelmingly that the per cent differences were unreliable The data for these responses were dropped from further consideration. Our criteria for dropping responses were as follows (*a*) in dealing with a single age group, an item was discarded if more than 90 per cent or less than 10 per cent of either the pubescent or the nonpubescent group checked the response in question, (*b*) in dealing with the combined age groups a response was discarded when in more than two of the nine subgroups 90 per cent or more, or 10 per cent or less of the subjects endorsed the response.

It will be observed that some of the subtests provide for only two possible responses to each item (*cf* Appendix, pages 62-70). In other tests there are three possible responses, *i.e*, an uncertain, or "?" response is included. These uncertain responses have not been included in the computations because of their scarcity. Hence, when in referring to the responses under consideration we refer to "left column" responses and "right column" responses we do not in any case indicate the uncertain (?) responses.

C. EVIDENCES THAT DIFFERENCES EXIST

Evidence that certain items of the questionnaire differentiated the premenarcheal and the post menarcheal groups was the first result obtained from the analysis of the data These data, given in terms of the critical ratios of the differences in the post menarcheal-premenarcheal response frequencies, are shown in the Appendix (pp 62-70).

Our next step was to inquire whether the differences obtained might be merely chance differences or whether they were true differences such as would arise only if the two groups of individuals really differed with respect to the functions covered by the questionnaire. Three lines of evidence bearing upon this question will be considered.

1. *Distribution of Critical Ratios*

The first evidence concerns the form of the distribution of the critical ratios of the differences in the proportions of the post menarcheal and premenarcheal groups making each response to each item, *i.e.,*

$$\frac{\text{Per cent of post menarcheal group making a given response} - \text{Per cent of premenarcheal group making the same response}}{\sigma_{diff}}$$

If the differences were due to randomly operating factors the distribution of critical ratios should approximate the normal frequency curve.

The data concerned are presented in Tables 7 and 8.[1] In these tables the percentage distributions of critical ratio frequencies are given for the following subject groupings: (*a*) the five age groups combined; (*b*) the 201 post menarcheal and 201 premenarcheal subjects equated for age, (*c*) the five age groups, individually considered. The distribution to be expected if there were no real differences in our groups with respect to the functions involved is also given. At the bottom of Tables 7 and 8 the constants of the distributions are given Chi square computed from the original frequencies before conversion to per cents and the probability that the differences between the obtained and the expected frequencies are true differences are given in the last two horizontal rows of the tables.

---

[1] In these and the following tables the number of responses varies from one distribution to another because of the variation in the number of responses which do not meet the percentage-frequency criteria for inclusion

TABLE 7

PERCENTAGE DISTRIBUTIONS OF THE CRITICAL RATIOS OF DIFFERENCES IN THE PROPORTIONS OF THE PREMENARCHEAL AND POST MENARCHEAL GROUPS MAKING LEFT COLUMN RESPONSES TO THE QUESTIONS,

CONSTANTS OF THE DISTRIBUTIONS, AND CHI SQUARE

A plus (+) critical ratio indicates a greater proportion of the post menarcheal group made the response

	Chance expec-tation	All sub-jects	201 sub-jects	Age group I	Age group II	Age group III	Age group IV	Age group V
—4 749- —4.250	00 00	00.00	00 00	00 00	00 00	00 32	00 00	00 00
—4 249- —3 750	00 01	00 31	00.00	00 00	00 00	00 00	00 00	00 00
—3 749- —3 250	00 05	00 31	01 16	00 00	00 00	00 00	00 00	00 00
—3 249- —2 750	00 24	01 26	01 16	00 30	00 00	00 32	00 32	00 95
—2 749- —2 250	00 92	03 77	01 45	01 51	01 20	01 62	01 29	00 95
—2.249- —1 750	02 78	08 18	08 12	03 32	03 00	03 90	01 93	04.76
—1 749- —1 250	06 56	07 86	10 14	09 97	06 61	07.79	09 32	06.98
—1 249- —0 750	12 10	09 75	12 46	13.29	10 51	15 91	14 47	11 11
—0 749- —0 250	17.47	11.32	12 75	15 11	12 91	19 48	19 94	14 29
—0 249- +0 250	19 74	18.87	15 07	13 29	16 82	15 58	17 68	15 24
+0 251- +0 750	17 47	09 12	10 14	17 52	20 42	12 99	20 26	18 41
+0 751- +1 250	12 10	08 80	09 86	11 18	15 62	12 99	07 07	12 70

TABLE 7 (Continued)

	Chance expectation	All subjects	201 subjects	Age group I	Age group II	Age group III	Age group IV	Age group V
+1 251- +1 750	06 56	05 97	06 96	10 27	07 81	04 22	04 18	05 08
+1 751- +2 250	02 78	07 23	04 64	03 32	03 90	03 25	01 61	05 40
+2 251- +2 750	00 92	04 72	02 90	00 60	01 20	00 97	01.29	01 90
+2 751- +3 250	00 24	00 63	02 61	00 30	00 00	00 32	00 64	00 63
+3 251- +3 750	00 05	01 26	00 58	00 00	00 00	00 00	00 00	00 95
+3 751- +4 250	00 01	00 63	00 00	00 00	00 00	00 32	00 00	00 32
+4 251- +4 750	00 00	00 00	00 00	00 00	00 00	00 00	00 00	00 00
+4 751- +5.250	00 00	00 00	00 00	00 00	00 00	00 00	00 00	00 32
No of responses	——	318	345	331	333	308	311	315
Mean critical ratio	00	— 02	— 10	— 03	+ 12	— 15	— 15	+ 10
σ_m	**	083	076	060	056	063	056	070
σ_{dis}	1 00	1 479	1 402	1 101	1 026	1 106	0 983	1 240
σ_σ	**	058	053	043	040	046	040	050
γ_1	0 00	+ 16	+ 19	— 04	+.21	— 03	+.25	— 32
σ_{γ_1}	**	14	13	13	13	13	.13	14
X^2*	—	237 36	225 89	22 81	13 51	18 92	21 41	62 37
P	—	000001	000001	.007	20	04	02	000001

*X^2 computed from the raw frequencies before conversion to percentages

**Varies with N

TABLE 8

PERCENTAGE DISTRIBUTIONS OF THE CRITICAL RATIOS OF DIFFERENCES IN THE PROPORTIONS OF THE PREMENARCHEAL AND POST MENARCHEAL GROUPS MAKING LEFT COLUMN RESPONSES TO THE QUESTIONS, CONSTANTS OF THE DISTRIBUTIONS, AND CHI SQUARE

A plus (+) critical ratio indicates a greater proportion of the post menarcheal group made the response

	Chance expectation	All subjects	201 subjects	Age group I	Age group II	Age group III	Age group IV	Age group V
—4.749- —4.250	00 00	00 00	00 00	00 29	00 00	00 00	00.00	00.00
—4 249- —3.750	00 01	00 88	00 28	00 00	00 00	00 30	00 00	00 29
—3 749- —3 250	00 05	00 88	00 28	00 29	00 00	00 30	00 00	00 00
—3 249- —2 750	00 24	00 88	01 66	00 29	00 00	00 30	00 89	00 86
—2.749- —2 250	00 92	02 35	01.38	01 16	01 15	01 19	01 18	01 44
—2.249- —1 750	02 78	06 47	04 97	05 80	01 44	03 27	04 44	02 87
—1 749- —1 250	06 56	06 76	06 63	07 54	08 62	07 44	07 99	04 60
—1.249- —0.750	12 10	09 12	08 29	14 78	14 66	11 61	10 36	07 47
—0 749- —0 250	17 47	10 29	14 36	15.65	16.09	12 20	18 05	08.33
—0 249- +0 250	19.74	19 41	15 19	17.97	16 67	17 56	18 93	16.09
+0 251- +0 750	17 47	10 29	11 60	13 33	17 82	20 83	17 16	18 97
+0 751- +1 250	12 10	09 12	11 33	10 14	12 07	11 90	12 15	15 52

TABLE 8 (Continued)

	Chance expectation	All subjects	201 subjects	Age group I	Age group II	Age group III	Age group IV	Age group V
+1.251- +1 750	06 56	10 88	10 77	08 41	06 32	06 55	07 10	10 06
+1 751- +2 250	02 78	06 76	08 56	03 48	02 87	04 46	01 18	06 03
+2 251- +2 750	00 92	03 82	01 66	00 87	02 01	01 49	00 59	05 17
+2 751- +3 250	00 24	01 47	01 93	00 00	00 29	00 30	00 00	01 15
+3 251- +3 750	00 05	00 59	00 55	00 00	00 00	00 30	00 00	01 15
+3 751- +4 250	00 01	00 00	00 28	00 00	00 00	00 00	00 00	00.00
+4 251- +4 750	00 00	00 00	00 28	00 00	00 00	00 00	00 00	00 00
+4 751- +5 250	00 00	00.00	00 00	00 00	00 00	00 00	00 00	00 00
No of responses	—	340	362	345	348	336	338	348
Mean crit. ratio	0 00	+ 08	— 19	— 13	+ 01	+ 05	+ 10	+ 41
σ_m	**	078	072	060	056	061	056	.068
σ_{dis}	1 00	1 433	1 378	1 123	1 031	1 122	1 018	1 256
σ_σ	**	055	051	043	039	044	039	048
γ_1	0 00	+ 17	— 05	— 15	+ 18	— 18	— 24	— 26
$\sigma\gamma_1$	**	13	13	13	.13	13	.13	13
*X^2	—	207 98	215 40	23 21	11 30	16 58	11 77	170 45
P	—	000001	000001	006	25	09	22	000001

*X^2 computed from the raw frequencies before conversion to percentages

**Varies with *N*

The crucial question upon which these data bear is as follows are the obtained differences greater than those which would be expected if the pubescent and nonpubescent groups did not differ in the frequency of their preferences for particular responses? Inspection of the percentage distributions of critical ratio frequencies for the five age groups combined indicates that the larger critical ratios occur with greater frequency and the smaller critical ratios with lesser frequency than would be expected if only random differences were operating The probability that these divergencies from the theoretical expectation might be due to chance has been tested in two ways with the following results: (*a*) according to the computed chi square such differences might arise by chance in less than one in a million times in the case of both the right and the left column responses, (*b*) the obtained standard deviations of critical ratio distributions of both the right and the left column responses are greater than expected by an amount approximately eight times their standard errors

These results, obtained from a consideration of all age groups combined, are confirmed by the matched groups of 201 pubescent and 201 non-pubescent subjects, and by the individual age groups, I-V inclusive. In the latter cases the reliabilities of the deviations of obtained from expected frequencies are not as great as for the combined age groups, but the directions are the same, except in the case of the left column responses of Age-Group IV. When the probability that five divergencies from the expected standard deviation as great as those occurring in the five age groups might arise by chance is calculated by Pearson's method (5), we find probabilities of less than .0003 and 00008 for the left and right column responses respectively.

The other characteristics of the distributions are not of great importance for the present problem. None of the distributions are significantly skewed, and the means of the distributions do not differ significantly from zero except in the case of the left hand response for Age-Group V We are unable to account for this shift.

2. *Consistency of Differentiation in Two Groups of Premenarcheal and Post Menarcheal Girls*

The second type of data which bear upon the main problem concerns the consistency with which *different groups* of pubescent and non-pubescent girls can be differentiated on the basis of their responses to the same questionnaire items. Data of this sort are available in the form of the items of the Pressey and Sullivan tests which were included in the present questionnaire These items were included upon the basis of the extent to which they differentiated between 175 post menarcheal and 175 premenarcheal girls in the schools of San José and Redwood City in the year 1934 (11) The location of these items in the present test follows:

Subtest I	Items,	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 17, 18
Subtest II	Items,	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 15, 17, 18, 20, 21, 23, 24, 27, 28, 30, 31, 32, 62
Subtest V	Items,	1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 13, 25, 26, 27, 28, 29, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 53, 54, 55, 56, 58, 59, 60, 62, 63, 64, 65, 66, 67, 68, 69, 71, 72, 73, 74, 75

The question may be raised as to whether these items differentiate the premenarcheal and post menarcheal girls of the present study as they did in the previous study In testing this possibility, responses which are favored by a greater proportion of premenarcheal girls are designated *premenarcheal responses,* and vice versa We have classified the responses as premenarcheal or post menarcheal as determined by the 1934 and the present data. These classifications are given in Table 9

TABLE 9

	Left Column Responses *Present study*		
1934 data	Premenarch responses	Post menarch responses	
Post m resp	11	18	X^2=22 43
Premenar resp	35	3	P=less than 00002
	Right Column Responses *Present study*		
1934 data	Premenarch responses	Post menarch responses	
Post m resp	3	37	X^2=23.62
Premenarch resp	18	11	P=less than 00001

The distribution of X^2 indicates clearly that the responses differentiate in the same general way on the two occasions although some responses are reversed We have not attempted to measure the exact extent of this relation inasmuch as the form of the question was changed in certain instances and the form of the answer in others

TABLE 10

Response			Item
Afraid	Not afraid		1 Smothering (IV-4), Suffocating (IV-9), Choking (IV-15)
"	"		2 Fire (IV-8), Flames (IV-23)
"	"		3 Going down a dark street (IV-10), Going into a dark room (IV-26)
Worry	Not worry		4 Death (IV-12), Dying (IV-28)
Afraid	Not afraid		5 Robbers (IV-31), Thieves (IV-24), Burglars (IV-13)
"	"		6 Crossing a high bridge (IV-19), Being on a high building (IV-56)
"	"		7 Children who start fights (IV-22), Bullies (IV-45)
"	"		8 Being followed on the street (IV-29), Some man chasing you (IV-66).
"	"		9 Bad dreams (IV-41), Nightmares (IV-62)
Yes	No	?	10. Do you feel awkward? (VI-54) Can you handle yourself well in walking and running? (VI-13).
"	"	"	11. Do you cry easily? (VI-17). Do you sometimes cry even when you try hard not to? (VI-68)
"	"	"	12 Are you allowed to use rouge and powder as much as you choose? (VI-27) Are you allowed to tint your fingernails as much as you choose? (VI-33)
"	"	"	13 Do you think boys are more capable school leaders than girls (VI-56) Should boys usually have the chief offices in school organizations? (VI-62).
			14 Have you secret ambitions and desires you seldom talk about? (VI-67). Do you imagine lots of things you never tell anyone? (VI-76) Do you have quite a few secrets you do not wish adults to know of? (VI-66)
			15 ( ) Have dates with boys... . or ( ) Go to girl's party? (V-17) ( ) Go to a party with a girl friend or ( ) Go to a party with a boy friend (V-24)
Like	Dislike	?	16 Run a day nursery (III-4) Care for small children (II-57)
"	"	"	17. Go to animal zoos (II-47) Go to a pet show (II-74)
"	"	"	18 Visit art galleries (II-45) Visit museums (II-46)

3 *Consistency of Responses to Items having a Similar Meaning*

The third sort of evidence as to the reality of the differences in the way premenarcheal and post menarcheal girls respond to the questionnaire items is the consistency with which responses of approximately the same meaning differentiate between the premenarcheal and post menarcheal girls. The items in Table 10 have been selected as having the most similar meanings. Subtest and item numbers are indicated in the parentheses

Of the foregoing items with similar meanings there are 16 pairs for which the percentage frequency criteria adopted by us (*cf.* p. 40) were met by the premenarcheal group In the case of 12 of these pairs of items (75%) the premenarcheal group favored[5] the responses that had the same meaning in the case of both items of each pair There were 18 pairs of items for which the percentage frequency criteria were met by the post menarcheal group In the case of 15 of these pairs (83%), the post menarcheal group favored responses having the same meaning in the case of both items of each pair These data indicate that there is a fair degree of consistency in the responses of our groups to items having similar or identical connotations This tendency is to be expected, of course, if the children are seriously considering each item and answering it according to the way it appeals to them

D NATURE OF THE DIFFERENTIATING ITEMS

While making up this questionnaire we had to use tentative hypotheses concerning regions or zones of interests and attitudes wherein premenarcheal-post menarcheal differences were likely to exist. With these in mind, items were formulated or were selected from the available test or questionnaire sources to tap these zones of interests and attitudes. For a consideration of the meanings of the differentiating responses, these items have been re-assembled and placed under the categorical headings used while making up the test

Naturally, any analysis based upon meaningful groupings of items is highly subjective because of a necessary degree of judgment one must exercise in deciding what items should be included in a given group and how the possible responses to the items should be interpreted. In the analysis which follows, we have tried, as far as possible, to avoid serious errors by taking only items whose responses seemed classifiable without equivocation. It is true, of course, that *none* of the items are indicative for a single region of interest ex-

---

[5] In this and the analyses that follow, the term *favored* is used to designate the responses which *are made by a greater proportion of the group*

TABLE 11

Subtest I

Item	6	Wrong	Not wrong*	?	Go to dances
"	12	"	" *	"	Indulge in petting or necking

Subtest II

Item	51	Like*	Dislike	?	Read romantic stories
"	62	" *	"	"	Flirt
"	64	" *	"	"	Go with a boy to a party or a dance
"	73	" *	"	"	Attend formal dances

Subtest III

Item	49.	" *	"	"	Be a housewife

Subtest V

Item	15	( )	attend a play	or	(*)	go to a dance?
"	17	(*)	have dates with boys	or	( )	go to a girl's party?
"	21	( )	tell and listen to stories	or	(*)	go to a dance?
"	24	( )	go to a party with a girl friend	or	(*)	go to a party with a boy friend?
"	36	( )	a candy factory	or	(*)	soldiers marching?
"	40.	(*)	a Hindu Prince	or	( )	a beautiful doll?
"	42	( )	*a bakery window*	*or*	(*)	*boys in uniform?*
"	48.	( )	roller skating	or	(*)	a dancing party?
"	49	( )	playing circus	or	(*)	the boys you have met?
"	50	( )	many boy friends	or.	(*)	one steady boy friend?

Subtest VI

Item	14	Yes*	No	?	Do you find it easy to make friends with boys?
"	32	"	" *	"	Do you get to dance as much as you like to?
"	46	" *	"	"	*Is there a particular boy you would like to have* for a friend?
"	53.	" *	"	"	Do your parents or teachers say you are too much interested in boys?
"	56	" *	"	"	Do you think that boys are more capable school leaders than girls?
"	60.	" *	"	"	Do you hate to meet a group of boys that you don't know very well?
"	62	" *	"	"	Should boys usually have the chief offices in school *organizations?*
"	71	" *	"	"	Would you like to have more boy friends?
"	78	"	" *	"	Would you like it better if boys and girls went to separate schools?
"	86	" *	"	"	Do your parents bother you by teasing you about your boy friends?

clusively; it is a matter of classifying the items into categories indicative of the dominant interest expressed by the response. Naturally it is to be regretted that some zones of interest have so few items that trends of differentiation can be indicated only in a very tentative way

1. The items of Table 11 are grouped under the heading "heterosexual interests and attitudes" An asterisk indicates the heterosexual response.

As stated above, the responses to the foregoing items have been rated as to their significance with respect to heterosexual interests. Thus for Item 6, Subtest I, "go to dances," the response "not wrong" was rated as indicating more heterosexual interest than "wrong." For Item 12, Subtest I, "indulge in petting or necking," the response "not wrong" was rated as indicative of stronger heterosexual interest than "wrong"

Of the responses for which the percentage frequencies were adequate to meet our criterion there were 21 that were rated as indicating greater heterosexual interest Nineteen of these were favored by a greater proportion of the post menarcheal than of the premenarcheal groups Of the 20 responses indicating less heterosexual interest, 18 were favored by a greater proportion of the premenarcheal girls These results suggest rather strongly that post menarcheal girls favor responses indicative of the stronger heterosexual interest.

2 The items of Table 12 are grouped under the heading "family adjustments" The asterisk indicates the response denoting adjustment, *those not marked presumably indicate friction in the family* situation.

For 35 of the foregoing items the per cent frequencies were adequate for further consideration Eighteen of the responses denoting adjustments were chosen in greater proportion by the premenarcheal girls, 17 were chosen in greater proportion by the post menarcheal girls This result was somewhat surprising and ran counter to our a priori judgment concerning this zone of attitudes and interests. Certainly the above data give no support to the hypothesis that post menarcheal girls give more evidence of open revolt from parental discipline and home restrictions than premenarcheal girls

3. The items of Table 13 were grouped under "adornment and

TABLE 12

Subtest VI

Item	16	Yes*	No	?	Does your mother understand you well enough that you feel like talking over your problems with her?
"	20	"	" *	"	Do you frequently disagree with grownups about what you should eat?
"	21	" *	"	"	Do your folks let you stay out at night as late as you would like?
"	22	"	" *	"	Do you usually have to go to bed earlier than you like?
"	23.	" *	"	"	Do you get to wear high heeled shoes as often as you wish?
"	24	" *	"	"	Do you have as many party or evening dresses as you think you need?
"	26	"	" *	"	Do you have to do more housework than you would like?
"	27	" *	"	"	Are you allowed to use rouge and powder as much as you choose?
"	28	" *	"	"	Do your folks let you decide how to wear your hair?
"	29.	" *	"	"	Are you allowed to have as many dates with boys as you would like?
"	30	"	" *	"	Do grownups object to your having boy friends?
"	31.	" *	"	"	Are you allowed to choose your own girl friends?
"	32	" *	"	"	Do you get to dance as much as you like to?
"	33	" *	"	"	Are you allowed to tint your finger nails just as you choose?
"	34	" *	"	"	Are you allowed to go to as many parties as you want to go to?
"	35	"	" *	"	Do you ever want to wear dresses which grownups think are too extreme?
"	47	"	" *	"	Do your parents seem to think you are a lot of trouble?
"	49.	"	" *	"	Do you think your parents are too strict with you?
"	50	"	" *	"	Do you ever have trouble with your parents because you want to smoke?
"	57.	"	" *	"	Would you like to go away from home to a boarding school?
"	61	"	" *	"	Is your life so humdrum or sad that you would like to get away where things are more interesting and exciting?
"	63	" *	"	"	Should girls of your age always obey their parents?
"	64.	" *	"	"	Are you satisfied with the amount of spending money given you?
"	65.	"	" *	"	Do you conceal lots of your worries and troubles from your folks?
"	66	"	" *	"	Have you quite a few secrets that you do not wish adults to know of?
"	69	"	" *	"	Do you think your folks baby you, or make too much over you?

TABLE 12 (Continued)

Item	73	Yes	No*	?	Do you sometimes feel that you are a burden to your family?
"	75	"	" *	"	Do you wish you had more freedom?
"	81	" *	"	"	Are you usually allowed to choose your own clothes when buying?
"	82	" *	"	"	Are you allowed to go riding with boys in day time when you wish?
"	83	"	" *	"	Do you hate to tell your parents of troubles you have had with other girls or with your teachers for fear they will interfere and make a fuss?
"	84.	" *	"	"	Are you frequently told to wash your hands or clean your finger nails?
"	85	"	" *	"	Do your parents find fault with your table manners quite often?
"	86	"	" *	"	Do your parents bother you by teasing you about your boy friends?
"	87	"	" *	"	Do your parents embarass you by bragging about you to other people?
"	88	" *	"	"	Are you allowed to use your spending money as you wish?
"	89	"	" *	"	Do your parents insist on taking you to parties, movies, etc, when you would rather go alone or with your friends?
"	90	"	" *	"	Do your parents or teachers object to the things you like to read?
"	91	"	" *	"	Do your folks think you don't spend enough of your free time at home?
"	93.	"	" *	"	Do your folks worry too much about your overdoing or hurting your health?
"	94	"	" *	"	Does it bother you if your parents want to know every single thing that happens at the parties you attend?
"	95	"	" *	"	Do your folks frequently object to the way you wear your clothes?

display of person." An asterisk marks the response indicating interest in adornment or display

The responses to the foregoing items have been rated as indicating *greater* or *less* interest in adornment and bodily display. Twenty-four of the above responses denoting greater interest had per cent frequencies that met our criterion Of these, 18 were favored by a greater proportion of the post menarcheal girls than of the premenarcheal girls The data indicating less interest are exactly the converse of the foregoing In this zone of interest, our working hypothesis that post menarcheal girls would have a stronger interest in adornment and display of person has received relatively strong support.

TABLE 13

Subtest II

Item	69	Like*	Dislike	?	Use cosmetics (rouge, powder, lipstick)
"	71	" *	"	"	Go to a fancy dress party
"	73	" *	"	"	Attend formal dances

Subtest III

Item	3	Like*	Dislike	?	Be a mixer of perfumes
"	5.	" *	"	"	Do aesthetic (artistic) dancing
"	6.	" *	"	"	Be a model or mannikin in fashion shows
"	14.	" *	"	"	Design jewelry
"	17	" *	"	"	Sell cosmetics (powder, rouge, perfume, etc.)
"	18	" *	"	"	Operate a beauty parlor
"	19	" *	"	"	Design fine gowns
"	23	" *	"	"	Sell Misses' clothes
"	28	" *	"	"	Design and make hats
"	54	" *	"	"	Be a circus performer
"	55.	" *	"	"	Be an artist's model
"	62	" *	"	"	Be a bathing beauty
"	63.	" *	"	"	Be a chorus girl

Subtest IV

Item	6	Worry*	Not worry	?	Personal appearance

Subtest V

Item	3	(*)	make silk underwear	or	( )	have stories read to you
"	9.	(*)	do hair dressing	or	( )	play circus?
"	18	( )	play with a pet kitten	or	(*)	do folk dancing?
"	38.	(*)	a fancy dress ball	or	( )	a baby nursery?
"	70.	( )	a box of paints	or	(*)	a party dress?
"	74.	(*)	Some party handkerchiefs	or	( )	lots of ice cream?

Subtest VI

Item	84.	Yes	No*	?	Are you frequently told to wash your hands or clean your finger nails?
"	85	"	" *	?	Do your parents find fault with your table manners quite often?

4. Items concerning games and activities of varying degrees of strenuousness are given in Table 14. The asterisk indicates the more strenuous response.

Again we have reported only those reponses for which the per cent frequencies are adequate. Of the 22 responses indicating greater interest in the more strenuous activities, 19 are favored by a greater proportion of the premenarcheal than of the post menarcheal group. Of the 21 responses indicating less interest in strenuous activities, 19 were favored by a greater proportion of the post menarcheal than of the premenarcheal group. These results strongly support the view

TABLE 14

Subtest II

Item	2	Like*	Dislike	?	Play on a teeter-totter
"	4	" *	"	"	Play Drop the Handkerchief
"	6	"	" *	"	Just loaf and be lazy
"	10	" *	"	"	Play basketball
"	12	" *	"	"	Play Follow the Leader
"	15	"	" *	"	Play Pussy Wants a Corner
"	17	" *	"	"	Play in the gym on the bars, rings, etc
"	18	" *	"	"	Hop and skip, or Hippity-hop
"	26	" *	"	"	Swing
"	30	" *	"	"	Jump rope
"	34	" *	"	"	Play Hide and Seek
"	35	" *	"	"	Just run and romp
"	39	" *	"	"	Play golf
"	40	" *	"	"	Ride horseback
"	55	" *	"	"	Be on an athletic team
"	59	" *	"	"	Climb cliffs

Subtest III

Item	5	Like*	Dislike	?	Do aesthetic (artistic) dancing
"	8.	" *	"	"	Conduct gym classes
"	20	" *	"	"	Be a playground director
"	54	" *	"	"	Be a circus performer
"	65	" *	"	"	Be a lifeguard at a pool or beach

Subtest V

Item	6	(*)	play Hop-Scotch	or	( )	hear the band play?
"	14	( )	drive an auto	or	(*)	roller skate?
"	18	( )	play with a pet kitten	or	(*)	do folk dancing?
"	30	(*)	a playground director	or	( )	a magazine writer?

that with the onset of puberty there is a change in interest with respect to the expenditure of energy in strenuous exercises

5 Items denoting fears, worries, and anxieties concerning possible injury or assault upon self or members of family are given in Table 15

The foregoing responses were rated as indicating more or less fear, worry, or anxiety. In the case of 24 responses indicating more fear, worry, or anxiety, 16 were given by a greater proportion of the premenarcheal girls. Of the 28 responses indicating less fear or anxiety, 18 were made by a greater proportion of the post menarcheal girls These data indicate a very slight tendency for the premenarcheal girls to indicate fears, worries, and anxiety responses more frequently than post menarcheal girls.

6 Items concerning fears and worries about health, death, and morbid subjects, are given in Table 16

TABLE 15

Subtest IV

Item	1.	Afraid	Not afraid	?	Collisions
"	2.	"	"	"	Murder
"	3	"	"	"	Poison
"	4	"	"	"	Smothering
"	8.	"	"	"	Fire
"	9	"	"	"	Suffocating
"	10.	"	"	"	Going down a dark street
"	13	"	"	"	Burglars
"	14.	"	"	"	Being alone in the house at night
"	15	"	"	"	Choking
"	18	"	"	"	Guns
"	22.	"	"	"	Children who start fights
"	23	"	"	"	Flames
"	24	"	"	"	Thieves
"	26.	"	"	"	Going into a dark room
"	29	"	"	"	Being followed on the street
"	30	"	"	"	Cyclones or tornadoes
"	31.	"	"	"	Robbers
"	43	"	"	"	Dogs
"	44	"	"	"	Playing rough games
"	45	"	"	"	Bullies
"	50	"	"	"	Thunder and lightning
"	51.	"	"	"	Learning to swim
"	55	"	"	"	Kidnappers
"	58	"	"	"	Drunken men
"	60	"	"	"	Spiders
"	63	"	"	"	An auto slowing down where you're walking at night
"	64.	"	"	"	Leaving your window open at night
"	66.	"	"	"	Some man chasing you
"	69	Worry	Not worry	"	Being punished

Subtest VI

Item	5.	Yes	No	?	Do you often worry about things that might happen to you or your family?

The responses to items in this group were rated according to their indication of more or less fear or worry. Of the 9 responses indicating more fear or worry, 4 were favored by a greater proportion of the premenarcheal girls. Of the 10 responses indicating less fear and worry, 6 were favored by a greater proportion of the premenarcheal girls. In this case there is no indication of a significant difference in the type of response favored by the post and by the premenarcheal girls of this study.

7. All items concerning fears, worries, and anxieties may be combined. Of the 50 responses which give an indication of relatively more fear, worry, or anxiety 30 are preferred by a greater pro-

portion of the premenarcheal girls Of the 61 responses indicating relatively less fear, worry, or anxiety 39 are favored by a greater proportion of the post menarcheal girls. Here we find a very slight tendency for the premenarcheal girls to indicate more fears than the post menarcheal girls

8 There are five items (Table 17) which refer to imaginative or secretive activities. An asterisk indicates the more imaginative or secretive response

In all of the foregoing instances a greater proportion of the post menarcheal girls indicate that they engage in or are interested in imaginative day dreaming activities. It is regrettable that more items of this type were not used in the test battery.

TABLE 16

					Subtest IV
Item	11	Afraid	Not afraid	?	Tuberculosis
"	12	"	"	"	Death
"	16	"	"	"	Health examinations
"	21.	Worry	Not worry	?	Feebleness
"	28	"	"	"	Dying
"	38.	"	"	"	Dieting
"	49	"	"	"	Going insane
"	57	"	"	"	Going blind
"	59	"	"	"	Going deaf
"	65.	"	"	"	Thinking someone in your family might die
"	67	Afraid	Not afraid	?	Operations
					Subtest VI
Item	72.	Yes	No	?	Are you sometimes worried about your health?

TABLE 17

					Subtest II
Item	*14.*	*Like**	*Dislike*	*?*	*Just imagine things, day dream*
					Subtest VI
Item	66	Yes*	No	?	Have you quite a few secrets that you do not wish adults to know of?
"	67.	" *	"	"	Have you secret ambitions and desires that you seldom talk about?
"	76	" *	"	"	Do you imagine lots of things that you never tell anyone?
"	79	" *	"	"	Do you sometimes pretend that you know things, so that people won't think you are very young and innocent?

E. Interest Maturity

The items of the Pressey and Sullivan tests which were included in the present questionnaire were standardized by their authors in terms of the increasing or decreasing proportion of children who made each response with increasing chronological age. A mature response is one which uniformly increases in frequency with the increasing chronological age of children, and vice versa. With the aid of this information we have made up the following tables which show the relation between mature and immature responses on the basis of Pressey's and Sullivan's findings, and the frequency with which the responses were favored by premenarcheal and post menarcheal girls in the present study. In these tables the heading, "Response favored by post menarcheal," means that a greater proportion of the post menarcheal girls in comparison with the premenarcheal girls made the response in question. The items from which these data are derived are identified on page 47.

It will be seen that there is a very clear tendency for the responses preferred by the chronologically mature, according to the Pressey and Sullivan ratings, to be preferred also by post menarcheal individuals. This point is striking in view of the fact that the post menarcheal and the premenarcheal subjects compared are of identical chronological ages.

These data confirm those reported by us previously (11) in indicating that the interests of post menarcheal girls are more similar to the interests of relatively older girls than are the interests of premenarcheal girls of the same chronological age.

TABLE 18

		Right Column Responses		
		Response favored by		
		Post menarcheal	Premenarcheal	
Response favored by	Older	38	3	$X^2=25.51$
	Younger	10	18	P=.000004
		Left Column Responses		
		Response favored by		
		Post menarcheal	Premenarcheal	
Response favored by	Older	17	11	$X^2=19.27$
	Younger	4	35	P=.00007

F LIMITATIONS AND CRITICAL COMMENTS

In the appropriate places we have repeatedly directed attention to the fact that our distributions are not complete with respect to the chronological ages herein considered. Selective factors associated with grade placement of the accelerated and the retarded children of the community made it impracticable for us to test all children of the ages chosen. To have done that would have necessitated testing some children in the elementary grades and some in the senior high schools where the social milieu would have been quite different from that existing for the junior high school group Because of the selective factors that have operated, our data cannot be generalized without due consideration of these limitations.

We should not claim that the foregoing results are free from the many limitations of the questionnaire type of inventories that are widely used now-a-days in personality and behavior studies It is quite possible that our results suffer from discrepancies between verbal reports to items of the questionnaire and what an individual's overt behavior would be in concrete life situations Discrepancies of this kind could arise from several causes, for example, (*a*) from concealment or distortion of true replies, (*b*) from the fact that impulses to action may not be consciously recognized, (*c*) because impulses may be too vague to be identified with a particular, specific item of the questionnaire, (*d*) because impulses may not be identified verbally in the same manner by adolescent girls as by the men and women who made up the tests (when this is the case responses may be wrongly interpreted), and (*e*) because the questionnaire may fail to repeat the substance of items with sufficient frequency to insure obtaining reliable indications of the subject's real attitudes and interests

Despite these possible limitations, we think that some such approach as herein made is a legitimate preliminary mode of inquiry. It seems worth while at the outset to accept the tentative hypothesis that, in the main, responses to items of a questionnaire are sufficiently indicative of overt behavior in life situations to reveal many of the general zones of attitudes and interests in which differences between special groups, for example, of premenarcheal and post menarcheal girls, are wont to lie. Ultimately, however, these differences should be validated by systematic field studies of overt behavior If later experimental data reveal discrepancies between responses to items of our questionnaire and overt behavior in life situations, these discrepancies themselves may be just as interesting and as significant for certain aspects of adolescent psychology as the fact of similarities between the two.

Regrettably, the methods employed in this study are not instructive as to what causal factors brought about the differences in responses by our premenarcheal and post menarcheal groups. It is generally known that certain physical changes associated with the menarche are to be accounted for in terms of specific tissue responsiveness to hormones of the anterior lobe of the pituitary and to hormones from the gonads. Perhaps certain behavioral changes as manifested in interests and attitudes also can be accredited to hormones from one or both of these glands In this connection, however, one must not overlook or underestimate the importance of extrinsic factors, particularly those of an ideational type which undoubtedly would play a more potent role in man than in lower animals lacking man's cultural background. It is not unlikely that self-evident changes in adolescent physique (accelerated growth, enlargement of the accessory sexual organs, hirsuites, etc) and the onset of the menses give rise in the adolescent to new interests, new aspirations, and new conception of freedom and restrictions imposed by society on persons who have passed that milestone in development which in all ages has separated the child from the young adult Finally, there may be differences in the social stimulation directed by adults to post menarcheal and premenarcheal girls If so, these social pressures could account for new types of interests, attitudes, or other mental residuals that show up in adolescent responses to items of a questionnaire.

The aim of the present study did not include an investigation of the practical significance of differences in attitudes and interests of groups of premenarcheal and post menarcheal girls In this connection, however, it is appropriate to recall that there was always a great amount of overlapping of the responses of the premenarcheal and the post menarcheal groups to individual items Certainly interests and attitudes as tapped by our questionnaire differentiate the post menarcheal and premenarcheal groups much less strikingly than do anthropometric measures during these adolescent years (10)

G SUMMARY AND CONCLUSIONS

Three kinds of evidence are presented in this report which indicate that premenarcheal and post menarcheal girls of the same chronological ages and social status respond differently to the items of an attitude-interest questionnaire with which they were tested

1. Relatively large differences in the proportions of premenarcheal and post menarcheal girls making particular responses occur with greater frequency than would be expected if the groups were, in fact, random selections from the same populations
2. Items which differentiated between premenarcheal and post menarcheal girls in a previous investigation with entirely different subjects differentiated, with but few exceptions, in the same way in the present study.
3. Clusters of items having similar meanings tended to differentiate between premenarcheal and post menarcheal girls in the same way.

The following statements concerning the meanings of the obtained differences in responses by the premenarcheal and post menarcheal groups seem to be warranted.

1. Post menarcheal girls favor the chronologically "mature" response more frequently than premenarcheal girls; *i e*, greater proportions of similar age-groups of post menarcheal than of premenarcheal girls give responses that are favored by increasing proportions of girls who are chronologically older
2. A greater proportion of post menarcheal than of premenarcheal girls of similar chronological ages favor responses indicating heterosexual interests
3. A greater proportion of post menarcheal than premenarcheal girls of the same chronological ages give responses indicating interest in adornment or in display of person
4. A greater proportion of post menarcheal than of premenarcheal girls of the same chronological ages give responses indicating disinterest in participation in games and activities requiring vigorous or strenuous activity.
5. A greater proportion of the post menarcheal girls indicate that they engage in or are interested in imaginative or day dreaming activities.
6. The evidence is against the hypothesis that post menarcheal

and premenarcheal girls differ in the extent to which they report the presence of family frictions or of revolt against family discipline

H. APPENDIX

Upon the following pages is a reduced facsimile reproduction of the *Attitude-Interest Questionnaire* On many of the items are recorded the critical ratios of the differences in the post menarcheal-premenarcheal response frequencies (*cf* p 39).[9] Those responses for which no critical ratios are given failed to meet the criterion adopted for inclusion in these computations (*cf.* p 40) A plus sign before the critical ratio indicates that a greater proportion of the post menarcheal than of the premenarcheal group favored the response A minus sign signifies the opposite of this. Of course, the numbers written on the items were not on the test forms submitted to the subjects.

---

[9] The critical ratios given are taken from the grouped data Thus $+0$ indicates the value falls between 0 and $+249$, etc

ATTITUDE-INTEREST QUESTIONNAIRE

NAME ____________________ SCHOOL ____________________

AGE __________ DATE OF BIRTH __________ GRADE __________

FATHER'S OCCUPATION ____________________ IS HE EMPLOYED NOW? __________

MOTHER'S OCCUPATION IF SHE IS A WAGE EARNER ____________________

DO YOU HAVE A TELEPHONE? __________ DO YOU TAKE A DAILY PAPER? __________

HOW MANY MAGAZINES ARE TAKEN IN YOUR HOME? (Check right answer)

None ____ 1 or 2 ____ 3 or more ____

HOW MANY PEOPLE LIVE IN YOUR HOME? __________

HOW MANY ROOMS ARE THERE IN YOUR HOME? __________

PURPOSE

Nearly every girl at your age begins to wonder what she wants to do when she grows up She must decide whether to go to college and for what life work to prepare In order to help girls decide such things, one has to know a great deal about their likes and dislikes, their fears and worries, what they think is right and wrong, etc. You are asked to aid us in obtaining this sort of information. If you will answer the questions in this blank thoughtfully and honestly, it will be possible for advisors, teachers, and parents to be of more help to girls in these matters No one who knows you (teachers, parents, etc.) will see your paper, so do not hesitate to answer the questions exactly as you think or feel

PART I

WHAT THINGS DO YOU THINK ARE WRONG?

Here is a list of things that some people think are wrong Read through this list If you think a thing is wrong to do, place a check (✓) before WRONG If you do not think it is wrong, place a check before NOT WRONG If you are not sure, place a check before ? Just mark as you think yourself, remember that some people think one thing is wrong, and some think another thing is wrong

Here is an example showing how to check the items

✓ Wrong	Not wrong	?	Slapping baby brother
Wrong	✓ Not wrong	?	*Reading a newspaper*

Now check so as to show whether you think it is wrong or is not wrong to do the following

1	Wrong	Not wrong	?	Talk back
2	Wrong -.25	Not wrong +.5	?	Smoke
3	Wrong -2.0	Not wrong +.2	?	Get angry
4	Wrong	Not wrong +2.25	?	Spend money
5	Wrong -2.0	Not wrong +0	?	Peddle
6	Wrong	Not wrong	?	Go to dances
7	Wrong	Not wrong	?	Joke
8	Wrong +0	Not wrong +0	?	Be fussy (finicky)
9	Wrong -.5	Not wrong .5	?	Pawn jewelry
10	Wrong	Not wrong	?	Play cards
11	Wrong -.75	Not wrong +2.0	?	Be divorced
12	Wrong -.5	Not wrong +.5	?	Indulge in petting or necking
13	Wrong -.0	Not wrong .0	?	Be afraid
14	Wrong	Not wrong	?	Play hookey
15	Wrong	Not wrong	?	Quarrel
16	Wrong +0	Not wrong 0	?	Go to war
17	Wrong	Not wrong	?	Grumble
18	Wrong -.75	Not wrong +2.0	?	Use slang
19	Wrong	Not wrong	?	Swipe fruit from orchards
20	Wrong -.75	Not wrong +2.5	?	Stay out late at night
21	Wrong -.5	Not wrong +.75	?	Eat lots of candy

End of PART ONE Go to next page

Prepared by Calvin P Stone and Roger G Barker,
Department of Psychology, Stanford University
for studies on adolescence

Printed only for
experimental use

PART II

WHAT THINGS WOULD YOU DO FOR FUN?

Some girls like to do one thing and some another thing when they wish to have a good time. Here is a list of things to do just for the fun of it. As you read down this list, make a check (✓) before the word LIKE if you would have fun doing the thing. Make a check before DISLIKE if you would not have fun doing the thing. If uncertain, make a check before ? Remember that some people like one thing and some another, so mark the items according to your own likes and dislikes.

1	- 25	0	?	Write poems
2	-2.0	+1.5		Play on a teeter-totter
3	-.25	+.25	?	Play Hide the Thimble
4	-1.50	+1.25	?	Play Drop the Handkerchief
5	-.25	+1.50	?	Attend lectures
6	+2.5	-2.75	?	Just loaf and be lazy
7	-.75	+1.50		String beads
8	-1.50	+0	?	Play London Bridge
9	+0	+.25	?	Play cards (bridge, whist)
10	-1.5	Dislike	?	Play Basketball
11	-.75	+1.75	?	Play Jacks
12	-1.75	+2.0	?	Play Follow the Leader
13	Like	Dislike	?	Take hikes
14	Like +3.0	Dislike -1.75	?	Just imagine things, day dream
15	-1.5	+1.75	?	Play Pussy Wants a Corner
16	Like -.25	Dislike -.5	?	Play Here I Come Where From?
17	Like -3.75	Dislike +3.0	?	Play in the gym on the bars, rings, etc
18	Like -1.25	Dislike +1.75	?	Hop and skip, or Hippity-hop
19	-.25	+.25	?	Paint with water colors
20	Like -2.0	Dislike +1.25	?	Sew, knit, or do fancy work
21	Like -0	Dislike +0	?	Pretend you are a movie actress
22	-.75	+.75	?	Play with dolls
23	Like -1.0	Dislike +1.75	?	Listen to stories being read
24	-2.25	Dislike	?	Draw with pencil or crayons
25	-1.25	+1.75	?	Play house
26	-2.5	+3.0	?	Play tag games
27	-2.0	+2.75	?	Swing
28	Like -0	Dislike +.25	?	Tell legends and fairy tales
29	-2.0	+1.80	?	Tell and guess riddles
30	-2.5	+2.0	?	Jump rope
31	-.75	+1.15	?	Play school
32	Like -.25	Dislike +.25	?	Cut out paper things with scissors
33	+.50	0	?	Whistle familiar tunes
34	-1.5	+1.75	?	Play Hide and Seek
35	-2.25	+3.0	?	Just run and romp
36	Like	Dislike	?	Swim
37	Like	Dislike	?	Play Tennis
38	Like	Dislike	?	Camp out
39	+0	-0	?	Play Golf
40	Like	Dislike	?	Ride horseback
41	+.25	+.25	?	Attend afternoon teas
42	+0	+.25	?	Do puzzles
43	Like +0	Dislike -0	?	Play in band or orchestra
44	Like -1.50	Dislike +1.25	?	Ride on a Merry-go-round
45	-0	+0	?	Visit art galleries
46	+0	Dislike	?	Visit museums
47	Like	Dislike	?	Go to animal zoos
48	Like +.75	Dislike -1.50	?	Attend band or orchestra concerts
49	Like	Dislike	?	Go to musical comedies
50	+2.0	+2.0	?	Read the society page
51	+1.0	+1.75	?	Read romantic stories
52	Like	Dislike	?	Read movie magazines
53	Like	Dislike	?	Go to picnics
54	Like -.5	Dislike -.25	?	Read the Sunday School paper
55	-2.0	+1.75	?	Be on an athletic team
56	Like +3.5	Dislike -3.50	?	Read The Ladies Home Journal
57	-1.5	+1.50	?	Care for small children
58	Like +.75	Dislike +0	?	Read The National Geographic Magazine
59	-2.75	+1.75	?	Climb cliffs
60	Like -.75	Dislike	?	Collect things (stamps, coins, etc)
61	Like	Dislike	?	Read the funnies
62	+1.75	-1.75	?	Flirt
63	Like	Dislike	?	Go for a boat ride
64	Like +2.25	Dislike -2.0	?	Go with a boy to a party or a dance
65	Like +1.0	Dislike -.75	?	Go to a club or a society meeting
66	+1.0	Dislike	?	Go window shopping
67	Like	+1.50	?	Smoke
68	+2.25	+2.0	?	Stay out late at night
69	Like +2.50	Dislike -.5	?	Use cosmetics (rouge, powder, lipstick)
70	Like	+1.50	?	Go to Sunday School
71	+2.0	-.75	?	Go to a fancy dress party
72	Like	Dislike	?	Go to football games
73	+1.5	+1.50	?	Attend formal dances
74	-3.5	+2.75	?	Go to a pet show

End of PART TWO   Go to next page

PART III

WHAT DO YOU WANT TO DO OR TO BE WHEN YOU GROW UP?

Here is a list of things that one might be or do Suppose that you could prepare yourself for any of them, and that they all paid about the same wage Carefully read each item on the list and place a check (√) before LIKE if you would like to be or do what that item states Place a check before DISLIKE if you would not like to do that kind of thing If you are uncertain, place a check before ? Remember that some people would like to do one thing and some another, so mark the items according to what you would like and dislike to do when you grow up

1	+.25	-.15	?	Be a private secretary
2	+1.0	-0	?	Conduct music classes
3	0	+.75	?	Be a mixer of perfumes
4	0	+.15	?	Run a day nursery
5	Like -.75	Dislike +1.0	?	Do aesthetic (artistic) dancing
6	Like +0	Dislike 0	?	Be a model or mannikin in fashion shows
7	-.75	0	?	Be a librarian
8	+1.50	+.25	?	Conduct gym classes
9	-0	+.75	?	Manufacture toys
10	Like -10	Dislike +.75	?	Work in the baby department of a big store
11	Like +0	Dislike 0	?	Plan menus for a big hotel
12	Like -0	Dislike +.75	?	Be head cook in a big hotel
13	-1.0	+.75	?	Be a waitress
14	+0	0	?	Design jewelry
15	Like +0	Dislike +.5	?	Do hand-craft work (weaving, pottery, etc)
16	+.25	-.25	?	Do window decorating
17	Like -0	Dislike -.75	?	Sell cosmetics (powder, rouge, perfume, etc)
18	+0	0	?	Operate a beauty parlor
19	+1.5	0	?	Design fine gowns
20	-.25	+.25	?	Be a playground director
21	Like +1.5	Dislike -1.25	?	Be a saleswoman in a book store
22	-2.0	+2.0	?	Work in a candy factory
23	+1.75	-2.25	?	Sell Misses' clothes
24	+2.25	-0	?	Sing over the radio
25	-1.0	+.75	?	Be a doctor (physician)
26	+1.0	0	?	Be a hospital nurse
27	-2.0	+2.0	?	Run a flower store
28	+1.0	.75	?	Design and make hats
29	+.75	.75	?	Conduct a cooking school
30	+.75	+.75		Design furniture
31	+2.25	-2.0	?	Plan radio programs
32	Like	Dislike	?	Give public lectures
33	Like	Dislike	?	Be a dentist
34	+.25	-0	?	Be a telephone operator
35	-.50	+0	?	Run a doll hospital
36	+2.0	-1.25	?	Write book reviews
37	-2.0	+1.50	?	Run a pet store
38	+1.50	-.25	?	Be a lawyer
39	Like -0	Dislike +.50	?	Be a typist or stenographer
40	+2.0	+1.75	?	Run a tea room
41	+0	+1.25	?	Be a postmistress
42	+1.0	1.0	?	Be an office manager
43	+0	+1.25	?	Be a Girl Scout Director
44	Like +.75	Dislike -1.75	?	Be a society reporter for a newspaper
45	-.50	+0	?	Be a missionary
46	+.50	0	?	Be a poetess
47	+1.75	-0	?	Be an opera singer
48	-.25	.50	?	Write for magazines
49	+0	+0	?	Be a housewife
50	-0	+0	?	Be an aviator
51	Like	+2.0	?	Be a magician
52	+1.25	1.25	?	Be a ladies' maid
53	+2.25	-1.50	?	Be a dean of girls
54	-1.50	+0	?	Be a circus performer
55	+.25	.75	?	Be an artist's model
56	-1.75	+1.25	?	Be an animal trainer
57	-0	+0	?	Be an actress
58	Like	+1.25	?	Be a science teacher
59	-.25	+.25		Be a sculptress
60	+.75	.50	?	Write novels
61	Like	+1.25	?	Be an English teacher
62	-.75	+0	?	Be a bathing beauty
63	-.75	+.75	?	Be a chorus girl
64	Like	Dislike	?	Be an evangelist (preacher)
65	-.75	+1.25	?	Be a lifeguard at a pool or beach
66	-.50	+0	?	Be a policewoman

End of PART THREE Go to next page

PART IV

FEARS AND WORRIES

Here is a list of things that people sometimes worry about, are fearful of, or are anxious about Read through this list If you are sometimes fearful of a thing or worry about it, place a check (√) before AFRAID or WORRY If you are not fearful of it or do not worry about it, place a check before NOT AFRAID or NOT WORRY If you are not sure, place a check before ? Remember that some people are afraid of one thing and some another, so mark the items according to your own feelings

1	Afraid	Not afraid	?	Collisions
2	Afraid	Not afraid	?	Murder
3	Afraid	Not afraid	?	Poison
4	Afraid	Not afraid		Smothering
5	Afraid -10	Not afraid + 25	?	Walking along the edge of a cliff
6	Worry	Not worry	?	Personal appearance
7	Afraid	Not afraid +.75	?	Rackets (terrible noises)
8	Afraid	Not afraid	?	Fire
9	Afraid	Not afraid	?	Suffocating
10	Afraid -2 5	Not afraid +2,5	?	Going down a dark street
11	Afraid	Not afraid	?	Tuberculosis
12	Worry	Not worry		Death
13	Afraid	Not afraid		Burglars
14	Afraid -10	Not afraid +1 25	?	Being alone in the house at night
15	Afraid	Not afraid	?	Choking
16	Worry	Not worry	?	Health examinations
17	Worry	Not worry	?	Wickedness (sin)
18	Afraid	Not afraid	?	Guns
19	Afraid	Not afraid -,50	?	Crossing a high bridge
20	Afraid	Not afraid	?	Danger
21	Worry	Not worry	?	Feebleness
22	Afraid	Not afraid +1 25	?	Children who start fights
23	Afraid	Not afraid	?	Flames
24	Afraid	Not afraid	?	Thieves
25	Afraid	Not afraid + 25	?	Looking down into deep water
26	Afraid	Not afraid +2.5	?	Going into a dark room
27	Worry	Not worry	?	Smoking
28	Worry	Not worry	?	Dying
29	Afraid -1.0	Not afraid +1.25	?	Being followed on the street
30	Afraid -10	Not afraid -1,5	?	Cyclones or tornadoes
31	Afraid	Not afraid	?	Robbers
32	Worry	Not worry	?	Work
33	Worry	Not worry	?	The end of the world
34	Afraid	Not afraid	?	Worms, caterpillars
35	Afraid	Not afraid	?	Graveyards at night
36	Afraid	Not afraid	?	Hell
37	Afraid	Not afraid	?	The sight of blood
38	Worry	Not worry	?	Dieting
39	Afraid	Not afraid	?	Strange noises
40	Afraid	Not afraid	?	Falling stars
41	Afraid	Not afraid	?	Bad dreams
42	Afraid	Not afraid	?	Mystery stories
43	Afraid	Not afraid	?	Dogs
44	Afraid	Not afraid	?	Playing rough games
45	Afraid	Not afraid	?	Bullies
46	Worry +75	Not worry -,25	?	People looking at you on the street
47	Afraid	Not afraid	?	Snakes
48	Afraid	Not afraid	?	The dark
49	Worry	Not worry	?	Going insane
50	Afraid	Not afraid +1 5	?	Thunder and lightning
51	Afraid	Not afraid	?	Learning to swim
52	Afraid	Not afraid	?	Noises at night
53	Afraid	Not afraid	?	Mice
54	Afraid	Not afraid	?	Dreams of drowning
55	Afraid	Not afraid	?	Kidnappers
56	Afraid +.75	Not afraid -,25	?	Being on top of a high building
57	Worry	Not worry	?	Going blind
58	Afraid	Not afraid	?	Drunken men
59	Worry	Not worry	?	Going deaf
60	Afraid	Not afraid	?	Spiders
61	Afraid	Not afraid	?	Earthquakes
62	Afraid	Not afraid	?	Nightmares
63	Afraid -1,5	Not afraid + 25	?	An auto slowing down where you're walking at night
64	Afraid	Not afraid	?	Leaving your window open at night
65	Worry + 25	Not worry -,75	?	Thinking someone in your family might die
66	Afraid -2.75	Not afraid +2 25	?	Some man chasing you
67	Afraid	Not afraid	?	Operations
68	Worry	Not worry	?	Being lost in a crowd
69	Worry	Not worry	?	Being punished

End of PART FOUR Go to next page

PART V

COMPARISONS

Here are some things to compare Read both things in each pair carefully Then make a check (✓) before the one thing in each pair that you prefer or like the better Here are two examples showing how to check the items

Which would you rather do

( ) wipe the dishes - - - - or - - (✓) go to the movies?
(✓) go on a picnic - - - - or - - ( ) clean house?

Now check the items below in the same way Remember to choose only the one in each pair that you would like the better

Which would you rather do

1 +(1.25) go to a ball game - - - - - - - - - - or -(1.5) wade in the water?
2 -(.5) play Jacks - - - - - - - - - - - - - - or +(1.5) drive an auto?
3 +(1.5) make silk underwear - - - - - - - - - or -(1.0) have stories read to you?
4 -(1.25) wade in the water - - - - - - - - - - or -(1.5) read the newspaper?
5 ( ) cut out pictures - - - - - - - - - - - or - ( ) ride in an auto?
6 -(1.75) play Hop-Scotch - - - - - - - - - - - or +(3.25) hear the band play?
7 ( ) read magazines - - - - - - - - - - - - or - ( ) play Punch and Judy?
8 -(2.25) ride a bicycle - - - - - - - - - - - - or +(2.0) play golf?
9 +(1.75) do hair dressing - - - - - - - - - - - or -(1.75) play circus?
10 ( ) cook a meal at camp - - - - - - - - - - or - ( ) dress paper dolls?
11 -(1.5) play circus - - - - - - - - - - - - - - or +(1.5) go to the opera?
12 +(2.0) read novels - - - - - - - - - - - - - - or -(2.0) ride on the merry go-round?
13 +(3.75) go to the movies - - - - - - - - - - - or -(3.75) visit the zoo?
14 +(1.25) drive an auto - - - - - - - - - - - - - or -(1.25) roller skate?
15 -(1.5) attend a play - - - - - - - - - - - - - or +(1.5) go to a dance?
16 ( ) dress up in old clothes - - - - - - - - or - ( ) watch athletic games?
17 +(1.5) have dates with boys - - - - - - - - or -(1.5) go to a girls' party?
18 -(1.25) play with a pet kitten - - - - - - - or +(1.25) do folk dancing?
19 +(1.25) go with your crowd or gang - - - - - or -(1.0) listen to the radio?
20 -(.5) do gym work - - - - - - - - - - - - - - or +(1.25) play Hop-Scotch?
21 -(1.75) tell and listen to stories - - - - - or +(1.75) go to a dance?
22 +(2.75) go with girls older than you - - - - or -(2.25) go with girls younger than you?
23 +(3.5) go with boys older than you - - - - - or - ( ) go with boys younger than you?
24 -(2.0) go to a party with a girl friend - or +(1.5) go to a party with a boy friend?

Which would you rather be

25 +(1.75) an architect - - - - - - - - - - - - - - or -(1.75) a cafeteria helper?
26 +(0) a cook - - - - - - - - - - - - - - - - - - or -(0) an actress?
27 -(1.25) a baker - - - - - - - - - - - - - - - - - or +(1.5) a judge?
28 -(2.0) a candy store clerk - - - - - - - - - - or +(2.25) a buyer for a store?
29 -(0) a lawyer - - - - - - - - - - - - - - - - - or +(0) a waitress?
30 -(2.0) a playground director - - - - - - - or +(2.75) a magazine writer?
31 -(1.0) Jane Addams, social worker - - - - - or +(0) Amelia Earhart, aviatrix?
32 +(1.75) Ethel Barrymore, actress - - - - - or -(1.0) Frances Perkins, Sec Labor?
33 +(2.0) Mrs Franklin D Roosevelt - - - - - or -(2.0) Helen Moody, tennis star?
34 +(2.0) Madame Schumann-Heink, singer - - - or -(2.0) Madame Curie, scientist?

Go on to next page

Which would you rather see

35 +1.0 pictures of famous women ------ or -1.0 a puppet show?
36 -2.75 a candy factory ---------- or +2.75 soldiers marching?
37 -1.25 animals in a zoo ---------- or +1.5 an auto race?
38 +.5 a fancy dress ball -------- or - 0) a baby nursery?
39 +1.25 an autumn scene ----------- or -1.25 a candy shop window?
40 +1.25 a Hindu Prince ----------- or -.75 a beautiful doll?
41 -.25 a lot of candy ----------- or +.5) a street accident?
42 -1.5 a bakery window ----------- or +2.0 boys in uniform?
43 -.5 a diving exhibition --------- or +.5 a class play?

Which would you rather think about

44 -.5 people in far-off countries ---- or +.5) college life?
45 ( ) having a palace ----------- or - ( ) getting ahead in life?
46 +2.5 visiting your best friend ----- or -2.5 pet animals?
47 -1.5 finding a lot of money ------- or +1.5 going on the stage?
48 +1.0 roller skating --- ------ or -.75 a dancing party?
49 -2.25 playing circus -- --------- or +2.5 the boys you have met?
50 ( ) having a palace ----------- or - ( ) being popular?

Which would you rather read

51 -1.75 My Dog Kiddo ----------- or +2.25 Treasure Island?
52 ( ) Little Women ------------ or - ( ) Peter Pan?
53 ( ) Green Lane Mystery --------- or - ( ) Billy Whiskers?
54 +2.0 Life of Theodore Roosevelt ----- or -2.0 Alice in Wonderland?
55 -.5 In the Days of the Giants ------ or +.25 Conrads Sea Stories?
56 +.5 The Legend of Sleepy Hollow ----- or -.5 Little Boy Lost?
57 -.0 Anne of Green Gables -------- or +.0) The Tale of Two Cities?
58 -1.5 The Ginger Bread Shop -------- or -+1.5 Daniel Boone?
59 ( ) The Golden Staircase -------- or - ( ) Hopsa, the Fairy?
60 ( ) Lorna Doone ------------ or - ( ) All About Peter Rabbit?
61 +1.75 Two Years Before the Mast ------ or -.75 Jack and the Bean Stalk?
62 +2.0 Uncle Tom's Cabin --------- or -2.0 Three Little Puppies?
63 -1.0 Tom Thumb ------------ or +1.0 Famous Movie Stars?
64 +2.5 Life of Benjamin Franklin ------ or -2.5 Billy Whiskers?
65 +1.25 A Night of Terror - -------- or -1.0 The Tale of Tom Tiddler?
66 -.25 Peter Rabbit ------------ or -+.5 Murder on Main Street?
67 ( ) Trust a Boy ------------- or - ( ) Six White Mice?
68 +2.25 Ivanhoe ---------------- or -2.25 Grimm's Fairy Tales?

Which would you rather have

69 +.75 A cedar chest ------------- or -.5) Lots of ice cream?
70 -1.5 A box of paints ------------ or +1.5 A party dress?
71 ( ) A picture story book -------- or - ( ) A new bedroom set?
72 ( ) A set of doll dishes -------- or - ( ) A set of golf clubs?
73 ( ) A croquet set ------------- or - ( ) A doll party?
74 +2.0 Some party handkerchiefs ------ or -2.25 Lots of ice cream?
75 -1.75 A doll trunk ------------ or +1.75 A bridge set?
76 +.75 Many girl friends ----------- or -.25 One or two close girl chums?
77 +.25 Many boy friends -- ------- or -.25 One steady boy friend?

End of PART FIVE Go on to next page

PART VI

QUESTIONS

Answer these questions according to what is true for you Answer yes by drawing a circle around (YES) answer no by drawing a circle around (NO) If you just can't decide or don't know, draw a circle around (?)

1 +Yes -.05 ? Do you usually feel well rested in the morning?
2 +.10 -.025 ? Do you think people find fault with you too much?
3 +.15 -.125 ? Do you often feel as if you had lost most of your pep?
4 +.25 -.10 ? Are your feelings frequently hurt so badly that you want to cry?
5 -.15 +.20 ? Do you often worry about things that might happen to you or your family?
6 -.175 +.125 ? Do you often have dreams that you don't like?
7 -.15 No ? Do you make friends easily?
8 -.25 -.075 ? Do you like to have people leave you alone or pay little attention to you?
9 Yes No ? Do you have dizzy spells quite often?
10 -.10 -.125 ? Do you feel lonesome a lot of the time?
11 +.125 -.15 ? Can you usually sit without fidgeting?
12 -.20 +.05 ? Do you giggle a good deal?
13 +.10 No ? Can you handle yourself well in walking and running?
14 +.25 -.15 ? Do you find it easy to make friends with boys?
15 +.10 -.15 ? Do your parents, forgetting that you are growing up, treat you as a child?
16 -.15 No ? Does your mother understand you well enough that you feel like talking over your problems with her?
17 -.15 +.075 ? Do you cry easily?
18 Yes +.225 ? Do you dream of being chased?
19 -.25 +.15 ? Do you, or would you like to keep a diary?
20 +.125 -.05 ? Do you frequently disagree with grownups about what you should eat?
21 Yes -.20 ? Do your folks let you stay out at night as late as you would like?
22 -.125 +.10 ? Do you usually have to go to bed earlier than you like?
23 +.175 -.075 ? Do you get to wear high heeled shoes as often as you wish?
24 +.225 +.175 ? Do you usually get to wear what you choose?
25 +.10 -.10 ? Do you have as many party or evening dresses as you think you need?
26 +.125 -.05 ? Do you have to do more housework than you would like?
27 +.40 -.075 ? Are you allowed to use rouge and powder as much as you choose?
28 +.20 +.15 ? Do your folks let you decide how to wear your hair?
29 +.25 No ? Are you allowed to have as many dates with boys as you would like?
30 -.10 +.20 ? Do grownups object to your having boy friends?
31 Yes No ? Are you allowed to choose your own girl friends?
32 -.10 +.10 ? Do you get to dance as much as you like to?
33 +.175 -.15 ? Are you allowed to tint your finger nails just as you choose?
34 +.075 +.10 ? Are you allowed to go to as many parties as you want to go to?
35 +.15 -.20 ? Do you ever want to wear dresses which grownups think are too extreme?
36 +.10 +.075 ? Do you find it hard to think of things to say when you are with one of the most popular girls in school?
37 +.10 -.10 ? Do you mind introducing strangers at a party?
38 +.125 +.10 ? Do you feel embarrassed when entering a big room after everyone else is seated?
39 -.10 +.10 ? Do you hate to stand up and give oral reports before a class
40 +.25 -.05 ? Would you feel shy about inviting the most popular girls in school to your party?
41 Yes No ? Are you troubled with shyness when shopping down town?
42 +.10 +.05 ? Are you sometimes afraid to speak up in class even if you know the answer?
43 -.10 +.075 ? Do you ever wish you were more popular?
44 -.10 -.125 ? Are you invited to places as often as other girls?

Go on to next page

QUESTIONS (continued)

45 Yes No ? Do you sometimes cross the street so you won't have to meet someone?
46 +.175 -.175 ? Is there some particular boy you would like to have for a friend?
47 Yes -.150 ? Do your parents seem to think you are a lot of trouble?
48 +.125 -.125 ? Do you blush easily?
49 Yes No ? Do you think your parents are too strict with you?
50 Yes No ? Do you ever have trouble with your parents because you want to smoke?
51 Yes No ? Do grownups treat you like a mere child?
52 Yes No ? Do you like to wear more jewelry than grownups think is becoming?
53 Yes No ? Do your parents or teachers say you are too much interested in boys?
54 .75 .05 ? Do you feel awkward?
55 -.75 No ? Do you wish that you might grow up faster?
56 +.25 .10 ? Do you think that boys are more capable school leaders than girls?
57 +.75 .025 ? Would you like to go away from home to a boarding school?
58 -.175 No ? Do you think that the proper place for women is in the home?
59 +.125 .025 ? Do you think boys should have more freedom outside the home than girls?
60 -.135 +.10 ? Do you hate to meet a group of boys that you don't know very well?
61 Yes No ? Is your life so humdrum or sad that you would like to get away where things are more interesting and exciting?
62 +.200 +.025 ? Should boys usually have the chief offices in school organizations?
63 -.200 No ? Should girls of your age always obey their parents?
64 -.15 No ? Are you satisfied with the amount of spending money given you?
65 +.100 .05 ? Do you conceal lots of your worries and troubles from your folks?
66 +.75 -.075 ? Have you quite a few secrets that you do not wish adults to know of?
67 +.35 -.35 ? Have you secret ambitions and desires that you seldom talk about?
68 Yes +.10 ? Do you sometimes cry even when you try hard not to?
69 Yes No ? Do you think your folks baby you, or make too much over you?
70 Yes -.15 ? Do you know what you want to be or to do when you grow up?
71 Yes -.15 ? Would you like to have more boy friends?
72 Yes -.175 ? Are you sometimes worried about your health?
73 +.125 No ? Do you sometimes feel that you are a burden to your family?
74 Yes No ? Do you think you would be homesick if you had to leave your home?
75 +.5 -.25 ? Do you wish you had more freedom?
76 +.275 +.025 ? Do you imagine lots of things that you never tell anyone?
77 +.225 -.25 ? Do you do quite a few things that your folks don't know about?
78 Yes +.10 ? Would you like it better if boys and girls went to separate schools?
79 +.15 +.05 ? Do you sometimes pretend that you know things, so that people won't think you are very young and innocent?
80 -.175 +.125 ? Would you like to stay just as old as you are now for a long time?
81 +.125 +.05 ? Are you usually allowed to choose your own clothes when buying?
82 +.125 -.175 Are you allowed to go riding with boys in the day time when you wish?
83 Yes No ? Do you hate to tell your parents of troubles you have had with other girls or with your teachers for fear they will interfere and make a fuss?
84 -.125 +.225 ? Are you frequently told to wash your hands or clean your finger nails?
85 -.10 +.05 ? Do your parents find fault with your table manners quite often?
86 +.20 +.10 ? Do your parents bother you by teasing you about your boy friends?
87 +.10 -.125 ? Do your parents embarrass you by bragging about you to other people?
88 +.15 +.05 ? Are you allowed to use your spending money as you wish?
89 Yes No ? Do your parents insist on taking you to parties, movies etc when you would rather go alone or with your friends?
90 Yes No ? Do your parents or teachers object to the things you like to read?
91 +.15 -.5 ? Do your folks think you don't spend enough of your free time at home?
92 Yes No ? Do your parents think out-of-style clothes are all right for you to wear?
93 Yes +.20 ? Do your folks worry too much about your overdoing or hurting your health?
94 +.20 +.35 ? Does it bother you if your parents want to know every single thing that happens at the parties you attend?
95 Yes No ? Do your folks frequently object to the way you wear your clothes?

Now look over each page to make sure that each item has been marked

REFERENCES

1. BARKER, R G, & STONE, C P Physical development in relation to menarcheal age in university women *Human Biol*, 1936, **8**, 198-222.
2. GOODENOUGH, F L The Kuhlmann-Binet tests for children of pre-school age A critical study and evaluation Univ Minnesota Inst Child Welfare Monog, Series, No 2 Minneapolis. Univ Minn Press. 1929 Pp 146
3. KELLEY, T L Statistical Method New York: Macmillan, 1923 Pp xi+390
4. MARSHALL, F H A Sexual periodicity and the causes which determine it. *Phil Trans Roy Soc. London, Series B*, 1936, **226**, 423-456
5. PEARSON, K Appendix to Dr Elderton's paper on "The Lanarkshire milk experiment" *Ann Eugen*, 1932, **5**, 337-338.
6. PRESSEY, S L, & PRESSEY, L C Interest-Attitude Tests New York Psychological Corporation, 1933.
7. ——— Development of the interest-attitude tests. *J Appl Psychol*, *1933*, **17**, *1-16*
8. PRIESEL, R, & WAGNER, R. Gesetzmassigkeiten im Auftreten der extragenitalen sekundaren Geschlechtsmerkmale bei Madchen. *Zeits f Konstitutionslehre*, 1931, **15**, 333-352
9. PRYOR, H B Certain physical and physiological aspects of adolescent development in girls *J Pediat*, 1936, **8**, 52-62
10. STONE, C P., & BARKER, R G On the relationship between menarcheal age and certain measurements of physique in girls of the ages 9 to 16 years *Human Biol*, 1937, **9**, 1-28
11. ——— Aspects of personality and intelligence in post menarcheal and premenarcheal girls of the same chronological ages *J Comp Psychol*, 1937, **23**, 439-455
12. SULLIVAN, C A scale for measuring developmental age in girls *Catholic Univ Amer Stud Psychol & Psychiat*, 1934, **3**, No 4 Pp vii+65.

Department of Psychology
Stanford University, California

The Journal of Genetic Psychology, 1939, 64, 73-84

THE APPROACH-WITHDRAWAL PATTERN IN THE SOCIAL BEHAVIOR OF YOUNG CHILDREN*

Iowa Child Welfare Research Station, State University of Iowa

MARY A WHITE AND HAROLD M WILLIAMS

A STATEMENT OF PROBLEM

The purpose of this investigation was to study the approach-withdrawal pattern of behavior of young children in social situations. The term approach will be used to denote patterns of behavior in which the child initiates, participates in, or seeks to continue any form of social contact with another child The word withdrawal indicates a "retiring" from social relationships These behavior patterns in which the child participates may be of many kinds, viz., physical, verbal, and the like. Each child may express these patterns by various attitudes, such as aggressiveness, dominance, friendliness, or cooperativeness. In this problem, however, we were not primarily concerned with these, but rather with the approach or withdrawal as such. Aggressiveness, for example, as well as friendliness may suggest an approach characteristic, while other patterns such as submissiveness may involve withdrawal

Little has been done on actually developing means of measuring approach and withdrawal. In fact, the concept itself has received little attention. The justification for its use here is primarily based on Williams' (1) factor analysis of Berne's social behavior patterns in young children. In this study two general factors, characterized by the author as approach-withdrawal and ascendance-submission, were found.

The essential problems of this study may be stated as follows first, did the children studied show any individual consistency of approach-withdrawal in their social relationships; second, to what extent were the children's contacts confined to a few other children, third, did their ranking change significantly during one year of school life?

---

*Accepted for publication by George D Stoddard of the Editorial Board, and received in the Editorial Office on March 21, 1938

B Method

In the analysis the "social contact" was made the basis Though this may, in complex social situations, be difficult of definition, the generally overt nature of the contacts of young children makes it fairly definitely observable In general, the "contact" is defined as a period during which a child's responses are primarily directed toward another child. The typical contact has an initial phase, when the child's center of attention shifts observedly to another child; a middle phase of interaction; and a closing phase, where the attention is directed elsewhere. The duration of such contacts may vary widely, but the phenomenon is not difficult to judge under proper conditions.

Each child was judged on the degree of reaction in his contacts and on the frequency of his contacts. The degree of contact was measured by a new device developed in this study. The criteria of degree of approach were (*a*) initiating a contact, (*b*) participation in the contact over a period of time, and (*c*) seeking actively to continue it. The following gradations of response were used as a basis

(*a*). The child initiates, participates in, and actively seeks to continue the contact.

(*b*). The child initiates, participates in, but does not actively seek to continue the contact.

(*c*). The child does not initiate, but participates in and seeks to continue the contact.

(*d*). The child does not initiate, but participates cooperatively, without seeking to continue the contact.

(*e*). The child does not initiate, but accepts contacts passively, without participation in or desire to continue them

(*f*) The child avoids contacts, he seeks actively to get away.

(*g*). The child observes the activities of others without initiating a contact, or without any other child seeking to initiate a contact.

(1). "*Initiates*" was defined as the active approach to another child, beginning a new contact, (2) "*seeks to continue activity*" was defined as persistent approach behavior until the child's goal was in some degree accomplished. This involved more than the initial approach, as for instance going up to another child and asking for a sand pail several times (3) "*Participates*" was defined as indulging in a "give and take" of whatever character it might be, such as playing with another child in the sand In category (*b*), *"The child initiates, participates in, but does not actively seek to continue the contact"* would involve, then, for example, the observation that the child definitely made an approach by going up to another child, participated in this contact as by asking for the sand pail, waited for the other child's response, but made no further advance toward the other one, so that the contact ceased In category (*g*), *"The child observes the activities of others without initiating a contact, or without any other child seeking to initiate a contact,"* the child actively observed the activities of others without entering into their play. This demanded more than a momentary glance. In the recording of the child's activities, judgment was based to a large extent upon the direction of the attention of the child under observation at that time

The contacts were recorded during 10 five-minute periods the first semester and 7 five-minute periods the second semester, the latter being reduced to a comparable basis by multiplying by 10/7 Each child was observed continuously for five minutes in rotation This made the basic analysis a time-sampling for both the frequency of contact and the degree of contact.

In tabulating the reactions, the letter corresponding to the degree of participation was recorded by the side of the child's name. Immediately beside it was placed a number symbol representing the name of the child contacted by the subject The symbols "*o*" and "*x*" were used for contacts with a group and the observer respectively

The value of the degree of approach-withdrawal in the categories was determined by the pooled opinions of 20 judges—staff members and advanced graduate students in child development The method of mean rank was used, each judge being asked to rank the categories in the order in which they displayed the approach-withdrawal characteristics under the following instructions:

> Kindly rank the following categories in order with respect to the degree of approach-withdrawal in social relations which they exhibit in your judgment, giving the highest rank to the approach quality

The mean rank of the judges' ratings is given in Table I The mean rank attained by each category was then assigned as its weight. While this method ignores certain aspects of scaling pro-

TABLE 1

Category	Mean
a	1.0
b	2.15
c	2.95
d	4.05
e	5.3
f	6.95
g	5.6

cedure, it is believed that it gives the relative weights of the categories sufficiently accurately for the present purpose. For convenience the values finally assigned to the various categories are given in Table 2.

TABLE 2

Category	Mean
a	1.0
b	2.2
c	3.0
d	4.0
e	5.3
f	7.0
g	5.6

The social situations for the sampling were periods of free play among children familiar with each other in a relatively stable school group. This period was either indoors or outdoors and was utilized exclusively in recording the child-to-child contacts. Thus the data involve a more or less random sample of the children's behavior in a representative situation, the free play period in school.

C. Subjects

The subjects were 53 kindergarten children in the public schools of Tulsa, Oklahoma. There were two sections, there being 23 in one group and 30 in the other. The age range was from 55 to 68 months.

The parents of the children in these districts were of moderate means, belonging to the business and professional classes.

D. Experimental Results

1. *Reliability*

Two types of reliability were studied, the reliability with respect to the child and the reliability of the observation These reliabilities were calculated on both the number of contacts and the degree value

The reliability of the number of contacts, calculated on the basis of alternate five-minute periods, was .62±.06 (Spearman-Brown, .77). The coefficient of reliability is rather low, indicating a need for a longer period of observation.

The reliability in degree of contacts was determined by correlating the mean value of the odd and even single contacts for each child This reliability proved to be .74±.04 (Spearman-Brown, .85). This coefficient of reliability is also low, though moderately high for this type of observation. The agreement between two recorders observing simultaneously was .93±.02 in the number of contacts The correlation between observers on the degree of contact was .92±.02 These correlations show that two observers can achieve a high degree of reliability in recording both the number and degree of contacts

2. *Group Results, Individual Differences, and Sex Differences*

The data on the number of contacts and mean degree of contacts for fall and spring are given for the group in Table 3. In this tabulation the value in degree of contact for each child was his own mean degree

TABLE 3

Contacts	Mean	Standard deviation	Range
		Fall	
Number	51.9	9.6	35 to 81
Degree	3.3	1.7	2.0 to 4.5
		Spring	
Number	62.6	10.5	41 to 88
Degree	3.2	1.7	2.0 to 4.7

Though the standard deviations are not large, the range shows that some children have more than twice as many contacts in a constant time as others. The mean degree of contact fell near the middle of the scale The range in degree of contact was also substantial On these bases, therefore, individual differences are substantial.

There was a negative relation between the total number of contacts and mean degree of contacts In the fall the correlation between number and degree was —.24±.09. In the spring the correlation was .01 of a point higher When it is remembered that the highest degree of approach received the lowest numerical rank, it will be seen that there is some, though not a high, tendency for approach behavior to be associated with high frequency of contact.

There were no apparent sex differences in either degree of number of contacts. The mean for boys in degree of contact for fall and spring was 3.2; the mean of girls in degree of contact was 3.2 and 3 1 respectively. The mean of total number of contacts for the boys in the fall and spring was 53 and 62, the mean of total number of contacts for the girls for fall and spring was 51 and 62

Individual variability in degree of approach-withdrawal is one of the more interesting phases of the problem For this analysis individual variabilities in terms of standard deviation were calculated for each child. The mean value for these variabilities was 1 6 for the fall and 1.6 for the spring. The corresponding standard deviations were 2 and 2 respectively. The outstanding finding in this analysis is the fact that the average of these individual variabilities was practically the same as the group standard deviation. In other words, the difference for each child from contact to contact was as great as the difference between children in their mean degree of contact. Compared with this value the differences between children in variability (standard deviation of individual variabilities) were very small

3 *Exclusiveness*

Exclusiveness was studied in terms of the distribution of the child's contacts with other children. The cumulative per cent of the total number of contacts that were with the five most frequently contacted children was made the basis Five seemed to be a reasonable limit to set on this type of analysis.

The cumulative means and variabilities in exclusiveness are shown in Table 4. As the tabulation shows, there is an increase in per cent, of course, with the inclusion of more children. The data are shown graphically in Figure 1. The outstanding characteristic of this curve is a smooth negative acceleration. A sharp break in the curve might be interpreted as suggesting that among these children there was a definite limit to the size of the exclusive group The mean values for these children showed no such break within the first five.

TABLE 4

Children most frequently contacted	Per cent of contacts Mean	Standard deviation	Range
		Fall	
1*	18	5.9	10 to 36
2**	30	7.3	19 to 54
3	40	8.2	25 to 71
4	49	9.0	31 to 82
5	56	9.9	43 to 73
		Spring	
1	19	5.4	9 to 33
2	32	6.4	16 to 46
3	42	8.0	23 to 60
4	51	9.0	30 to 75
5	58	9.7	34 to 90

*Per cent of total number of contacts which were with child most frequently contacted.

**Per cent with two children most frequently contacted.

Exclusiveness can also be studied on a correlational basis. In this type of analysis the question is raised, "To what extent does the child who ranks highest in concentration of contacts with one other child also rank highest in concentration of contacts with two, three, or four children?" The correlation between number of contacts with child No. 1 (most frequently contacted) and child No. 2 was .26± .04, between child No. 1 and children Nos. 2 and 3 it was .03± .09, between children Nos. 1 and 2 and children Nos. 3 and 4 it was .25± .09. The tendency toward the formation of small cliques or groups was low, therefore, in these children in this social setting.

4. *Changes from Fall to Spring*

As Table 3 shows, there was an increase in the total number of contacts for each child in the spring. The critical ratio was 5.5, showing the difference to be highly significant. This suggests a growth in social behavior, for the group as a whole at least with respect to frequency of contact.

The correlation between the fall and spring total number of contacts was .21± .09. This low correlation denotes a substantial shift in rank in number of contacts among the children from the fall to the spring. Nine children increased 20 or more contacts in the spring, while 11 decreased from 7 to 19 contacts. Two decreased

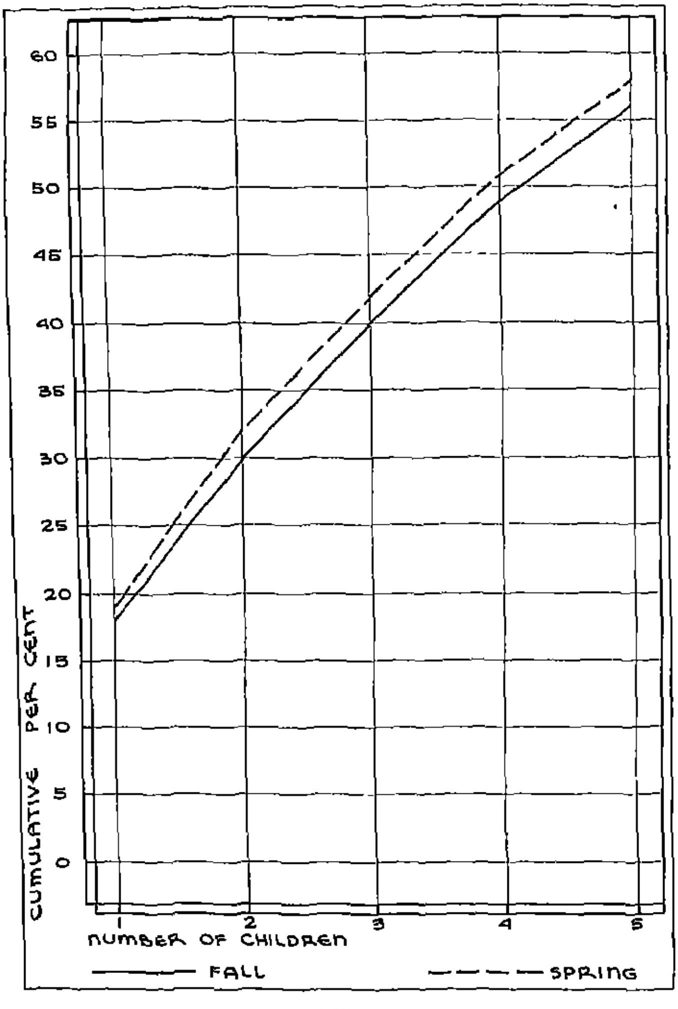


FIGURE 1
Cumulative Percentages for Varying Numbers of Children Controlled

by more than 30 contacts. Group variability did not change significantly during the year.

Practically no prediction was possible in the total number of different children contacted by each child in the spring from the number of children contacted in the fall, the coefficient of correlation being .11±.09.

The mean degree of contact for the group did not change from fall to spring. The fall-spring correlation was .50±.07. This is higher than the correlation between the number of contacts in the fall and spring, there being less shift in position during the year in degree than in number of contacts.

There was a slight tendency in the direction of a regression toward the mean in degree of contacts over the interval between observations. Nineteen children who had a mean score of 3.0 or below ascended the scale in approach, while 11 who had a mean degree of contact of 3.0 or below descended the scale toward withdrawal in the spring. Thirteen who had a score of 2.9 or above in the fall descended the scale, and only three who had a mean score of 2.9 or above were still ascending the scale in the spring. In some extreme cases this tendency toward a presumably desirable medium did not hold, as three of the four who had a score of 4.0 or below were still lower on the scale in the spring.

The correlation between the number of contacts with the child most frequently contacted in the fall and the spring was .16±.09. The child may change very much in exclusiveness, therefore, from fall to spring as measured by this criterion.

The shift in exclusiveness may be studied more analytically in terms of changes in status of the "best friends" (children most frequently contacted) from fall to spring. This question was analyzed by tabulating the frequencies with which the child who was most frequently contacted by a given child in the fall remained the most frequently contacted in the spring. This was carried on cumulatively for the first four choices. The results are given in Table 5. Thus of 295 instances only 82 or 27.7 per cent of the four children most

TABLE 5

	Spring rank					
Fall rank	First	Second	Third	Fourth	Other	Total
First	8	12	9	2	42	73
Second	12	7	6	0	65	90
Third	6	5	9	2	85	107
Fourth	0	1	1	2	21	25
Total	26	25	25	6	213	295

frequently contacted in the fall remained among the four most frequently contacted in the spring[1]

5 *Discussion*

The outstanding finding of this study is the great lability displayed by the children, whether expressed as individual variability over a short period of time, as exclusiveness, or as changes over a longer period, an individual variability equal to the variability of the group, a 72 per cent change in first choices or a drop of 40 points in reliability in a measure over an interval of less than nine months is of some significance for the problem of the stability of "personality traits" in young children.

The writers wish to take this opportunity of re-emphasizing the significance of the time-sampling technique where judgments can be made in terms of single instances of a behavior phenomenon, such as a "social contact," a "temper tantrum" or "acceptance of a suggestion" Under such conditions the technique can yield results comparable to rating techniques with this fundamental advantage, that the rating is not derived from a general impression, with all its possibilities of error from "halo" effect and unconscious overemphasis on single dramatic events. The time-sampling represents instead the mean or cumulative resultant of a series of such judgments, in which normal forgetting becomes a help rather than a hindrance, tending to force the judgment to be made principally in terms of the observation of the moment Since the judgment is made a matter of record at that time, it need no longer be carried in mind for the purpose of achieving some final "estimate" which is based in all probability in the usual rating technique on some unsystematically struck average of such observations.

The time-sampling technique has the further basic advantage of studying directly the desired phenomenon under any circumstances where incidence may be important. This may range from the incidence of the behavior under ordinary "real life" conditions, such as the frequency of appearance of "temper tantrums," to its incidence under specific or experimentally controlled circumstances, such as the incidence of attempts at ascendant behavior of one child paired with another in the experimental room

---

[1] Certain arbitrary decisions had to be made in this analysis All children attaining identical frequencies were given the same rank. On this basis, four grades of rank were not always possible Thus, for instance, if the first four children all received the same rank, all were counted as first choices and the analysis went no further Similarly, if there was one child in the first rank, two in the second, and four in the third, all seven were counted.

The method can have, therefore, an inherent validity in that it may utilize observations of actual behavior under the desired stimulating conditions. Under the proper circumstances its validation reduces itself to the checking of the original single observations and the problem of achieving an adequate representative sample It is also very flexible in its possibilities of adaptation to particular problems.

In connection with one result of this study, the variability in degree of approach exhibited by the children, it seems desirable to re-emphasize the importance of an accurate analysis and description of the stimulating conditions in most responses in the categories of social and personality behavior. It may truly be said of many researches to date that the child has been more subtle than the experimenter, responding differentially to nuances in the stimulus that were assumed tacitly to be irrelevant by the experimenter. In estimating a child's "obedience," for example, it may be crucially important to know by whom and how obedience is demanded In social reactions, the importance of the other person as a factor in the individual's behavior deserves far greater emphasis

E. Summary and Conclusions

In this investigation a study was made of the approach-withdrawal pattern of social behavior of 53 children by the observational method, using both frequency of contact and a scale of gradations in approach and withdrawal as criteria. In the scale the mean rank by 20 judges was the basis on which scale values were assigned. Observations were made in the fall and in the spring

There was a high degree of reliability between the two observers in recording both the number and the degree of contacts. The reliability of the scale in degree of contacts was higher than the reliability of the number of contacts, though both were fairly reliable for a relatively brief observation time.

1. *Group Results*

a. For the group there was a mean of 51.9 contacts in the fall and 62.6 in the spring in a 50-minute observation period. The total range was large, the number of contacts being twice as great for some children as for others

b. The mean degree of contact, fall and spring, fell near the middle of the scale. The standard deviations were small but the total range fairly large.

c. There was a correlation of —.24± .09 between the frequency of contact and the degree of contact, indicating a slight tendency for the approach characteristic and the number of contacts to be associated.

2. *Exclusiveness*

a. Exclusiveness for the group was fairly low, averaging 18 per cent with the most frequently contacted child for the fall and 19 per cent for the spring. The standard deviation was 5.9 per cent in the fall and 5.4 per cent in the spring, although the range extended from 10 to 36 per cent in the fall and from 9 to 33 per cent in the spring.

b. For the five most frequently contacted children, there was a slight negative acceleration in number of contacts. There was no sharp break in the curve for the group as a whole.

3. *Changes from Fall to Spring*

a. There was a low correlation in the total number of contacts for each child between fall and spring, which denotes a substantial change among individual children. Group variability remained about the same throughout the year.

b. There was a substantial growth of behavior with respect to frequency but very little with respect to degree of contacts. There was a slight tendency for the children to become more balanced in degree of contact, those at either extreme regressing toward the mean.

c. There was a decided change in individual exclusiveness through the year in that the percentage of identity in the four children most frequently contacted was only 27.7.

REFERENCE

1. WILLIAMS, H. M. A factor analysis of Berne's social behavior in young children. *J. Exper. Educ.*, 1935, 4, 142-146.

1123 South St. Louis Street
Tulsa, Oklahoma

Iowa Child Welfare Research S[illegible]
State University of [illegible]
Iowa City, [illegible]

The Journal of Genetic Psychology, 1939, 54, 85-105

A CRITICAL ANALYSIS OF CERTAIN PERFORMANCE TESTS*[1]

Department of Psychology, New York University

CHARLES M. MORRIS

A. HISTORICAL

Mental measurement, and intelligence tests in particular seem to have developed through four stages. First was the long "period of invention" when the chief efforts of workers in the field were devoted to devising tests and test items. Most were measures of the so-called "simpler mental processes or abilities" and included such tests as sensory acuity, mental imagery, reaction time, free and controlled association and memory problems. Usually a test was but a single item and was thought to measure a single "mental ability."

Secondly came a period of evaluation wherein the efforts of a few workers were devoted to empirical comparisons of the relationships between test scores and some criterion of mental ability such as school marks, estimates of intelligence, and the like. Included among these are the studies of Bolton, Boas, Gilbert, Wissler, Sharp and Titchener, and others. Of course the results obtained indicated little or no relationship.

The third period was ushered in by Binet's monumental theory of intelligence measurement by combining a relatively large number of "simple" tests into a scale which yielded a total intelligence score and culminated in his 1905 scale and its revisions. Thus he considered intelligence to be a "complex organization" rather than a number of independent "simple" mental traits and he proposed to measure it by sampling a wide variety of responses. Subsequent progress in statistics made available techniques for determining better combinations of items, thus obtaining closer relationships to the other criteria of intelligence. This period includes the development of the multitude of verbal group tests of intelligence as well as the numerous revisions of the Binet Scale.

---

*Received in the Editorial Office on March 23, 1938.

[1] This paper is based upon a Ph.D. dissertation done at New York University under the direction of Dr. Edwin R. Henry.

But while developments in statistics were providing tools for making better tests they were also making available tools for critical analysis of the already existing tests as well as the new ones Thus, the fourth period is one devoted to experimental and statistical studies prosecuted for the purpose of determining something about the nature of intelligence and what tests of intelligence were measuring Of course this period had its origin in the early work of Spearman but attained its full stature after the developments of Kelley, Hotelling, Thurstone, Tryon, and others

Since it has a direct bearing on the study to be reported here Thurstone's recent analysis (11) of the nature of mental ability is described.

He assembled a battery of 56 paper and pencil tests selected so as to give an adequate sample of all the different kinds of tests in existence and to include measures of many supposed "kinds" or "characteristics" of intelligence. The battery was administered to 240 college students and scores for each test were obtained All of the possible intercorrelations among the tests were then computed and the correlational matrix was analyzed by the centroid method of factorial analysis described by Thurstone (12) It was found that performance on the 56 tests could be accounted for on the basis of *seven* mental abilities or characteristics of intelligence That intelligence, conceived of as the total ability required to perform on all the tests, is not a general ability was shown by the fact that not one of the seven abilities was found to be involved in all the tests. That the seven abilities were independent primary abilities was demonstrated by the fact that, when conceived of as dimensions of hyperspace, they were orthogonal. From a psychological analysis of the tests in which each of the abilities was found Thurstone ventured an opinion on their nature and names as follows

> *Number Facility* "This primary factor is entirely restricted to numerical thinking and it is present in the highest amount in simple numerical speed tests It is less conspicuous in those numerical tests which involve reasoning or formulation of a problem in quantitative terms"
>
> *Word Fluency* "This ability is prominent in those tests in which the subject is asked to supply words in given context It is of considerable psychological interest that all the tests that signify this primary ability are limited to the recall of words and that none of the tests high in this factor involve sustained verbal reasoning"

> *Visualizing* "As far as can be determined at present this factor includes the visualizing of solid objects as well as flat space"
>
> *Memory* "The conclusion seems warranted that a person can be described as having a good memory in general without specification as to what he can remember well. However, the memory factor must be examined by experimental studies devoted to this particular field before its generality becomes a certainty This much is now known, however, that memory is distinct from other mental abilities"
>
> *Perceptual Speed.* "It is prominent in those tests in which the subject is asked to identify something quickly when it is mixed with other perceptual material."
>
> *Induction* "It is involved in several tasks which have in common the characteristic that the subject must discover some principle or rule that governs the material One of these tests was devised by Spearman as a test of the general intellective factor which he denoted '*g*' This primary trait is close to the general intellective ability which has been postulated by Spearman, and it may eventually turn out to be essentially the same factor The single-factor hypothesis of Spearman makes the assumption that all activities that may be called intellectual involve this central factor, but such is not our finding The ability to name opposites and synonyms is generally regarded as an index of general intelligence, but these activities have no inductive component in our analysis"
>
> *Verbal Reasoning* "It is exemplified by tests of verbal analogies and tests in which the subject is asked to match proverbs which have the same moral or quotations which have the same meaning and to make numerical estimates which require deductive reasoning" (11)

The next step for mental measurement seems to be obvious. If all of the now existent paper and pencil tests of intelligence numbering in the hundreds are measuring only seven characteristics it would seem wise to construct seven new tests each of which is a perfect measure of one and only one characteristic Thus, a battery of the seven tests would serve the same purpose now performed by the much larger number of tests

Although Thurstone's study is the only one reviewed in detail here the contributions of Spearman, leading to his two-factor theory, Thorndike and his multi-factor theory, Thomson's theory, Kelley, Hotelling, Holzinger, Tryon, and others and their resulting group factor theories should not be overlooked

All but Thorndike have developed statistical techniques for analyzing a correlation matrix into factors, but in some cases at least the number of factors obtained have been determined as much by the techniques as by the tests and measures used However, Thurstone's center of gravity technique is the general case of which all the others are but special cases. It has the advantage also of being free of many of the limiting assumptions of the others, and the very complete rational basis in the mathematics of the matrix theory can be made to lead to significant, meaningful variables (5).

While much concern has been evidenced with regard to the nature of paper and pencil tests, and especially those of the "verbal type", the case for performance tests is quite different Efforts in this field have been devoted to the construction and standardization of more and more tests quite like the first period alluded to above. True it is that some scales have been developed, notably the Pintner-Paterson, Arthur, Merrill-Palmer, Army-Performance, and Cornell-Coxe, but performance tests are still largely used as single individual measures of a single mental ability although it is claimed for some of them that they measure a complex or even a mixture of characteristics of intelligence Some studies have been conducted to determine the relationships between performance intelligence (here used to distinguish the results of performance testing from intelligence as measured by paper and pencil and verbal tests) and such criteria as were used to validate other tests, but the favorite method of validation still seems to be a correlation with the Stanford-Binet or some other such test. Seldom has an attempt been made to verify the assumption that a test measured what the author set out to measure, or what the name of the test indicates it is supposed to measure or even, for that matter, how many different things it might actually be measuring

Thus we have been claiming to measure, by performance tests, a complex general mental ability without bothering to determine whether or not there was a complex or if a complex were found, its nature, number of elements, etc. Murphy (7) in a study of the relation between mechanical ability tests and verbal and non-verbal intelligence tests has called attention to the confusion of terminology in the field of mechanical aptitude measurement. Her study made use of the factor analysis technique and disclosed the fact that the "trait which has been called mechanical ability is not a unitary trait, but rather, a complex" of a trait which she has called "speed of hand and eye coordination" and a trait "dependent on the mental manipulations of spatial relations." She also showed that her findings could not have been anticipated from the names by which the tests have been known; for the non-verbal type of intelligence tests had more in common with the mechanical aptitude tests than they did with the verbal intelligence tests

Some light is thrown upon the analysis of performance test abilities by an introspective study made by Gaw (3). She had seven psychologists take a large battery of performance tests, most of which are included in the present investigation, and report the introspections of what constituted their behavior in the test situation They reported that the tests measured principally reasoning as represented by the Picture Completion test, memory, of the type found in planning ahead the Porteus Maze solutions in relation to the preceding stages, and imagery of kinaesthetic and visual types Kinaesthetic imagery was reported as helpful in 10 of the 14 tests In addition to these factors which she would consider expressions of general intelligence, "a special ability, manual dexterity, was also measured although probably in only one test, the Goddard Form Board." To this list she adds the special ability to be found in the Dearborn Form Board, namely, perception of form relations, and the temperamental qualities of the Porteus Maze: "ability to take forethought, to plan ahead, and to work with deliberation."

Other studies less extensive than those described above have contributed to the variety of our knowledge of the factors involved Paschal (8) concludes from his standardization of the Witmer Cylinder Test that it calls for form and space perception and measures mechanical ability "Distributive attention may be brought out better in this test than in any other so far standardized," he believes Worthington (16) suggested that the Picture Completion Test, the Cancellation of *A*, and the Knox Cube are tests of special ability and "further analyses of these tests should be made to determine what ability they do test" The Maxfield Color Cube Test was first used by its author in connection with a study of image ability of children In addition to this ability, Hutt (6) considered the specific element of "mental noting," in which occur attention, perception, and an imprinting upon memory of an image of the two colored spatial pattern. This, however, may be little more than that which has been termed the memory span. Pintner and Anderson (9) reported that it is possible through administering the Picture Completion test "to get some measure of the child's apperceptive ability, to see how well he is able to meet the requirements of each situation by grasping its meaning" Goddard's Adaptation Board (4) was an attempt to "discover the ability of children to adapt themselves to changed conditions." Woolley and Fischer (15) considered the card sorting test a good measure of an eye-hand coordination And so we might continue to cite more studies and more alleged factors or mental abilities or characteristics of intelligence

In summary we see that the mental ability or abilities measured by performance tests of intelligence have not been satisfactorily determined even to the extent of their number, to say nothing of their names and descriptions. That many such abilities are supposedly being measured is evident from the variety of names and descriptions given by test builders and users on the basis of a priori assumptions or observations of the tests in clinical usage. It is the general purpose of the study to be reported here to attempt to produce data bearing on this problem and perhaps to solve some of the problems raised.

B. Problem

On the basis of the preceding survey and discussion a much-needed, general, long-term problem presents itself The ultimate results to be obtained should lead to a restatement of the qualities measured by performance tests and possibly to a reformulation of the nature of "performance intelligence".

The hypothesis formulated as the basis for the study runs somewhat as follows: A great many tests have been devised, all of which are claimed to be measures of general intelligence or some attribute or characteristic of it. Since these tests are assumed to be measures of the same thing, then large positive correlations should obtain among them Experience has indicated that large positive correlations are not found, therefore the tests must contain uniqueness (part of which is probably unreliability) The common elements of the intercorrelations, and hence of the tests, may be determined statistically, both as to the number and magnitude It seems likely that the common elements among the tests will not exceed five or six If such is the case then new tests may be constructed to measure the number of the common elements and the half-dozen new tests should provide all the information about a testee now obtained by the hundred odd tests in use Such new tests may be validated by correlating performance on them with the factor loadings of the appropriate common elements and should result in increasing the value and validity of performance measurement.

Specifically then the problems of this investigation may be stated as follows

1. What is the nature of the interrelationships found to obtain among performance tests of intelligence?

2. Is there a general intelligence measured by performance tests?

3 Are there group factors or traits measured by performance tests?

4. How many abilities, characteristics, or kinds of intelligence are measured by performance tests?

5. Are there unique performance abilities, i.e., are there tests which have nothing in common with other performance tests?

6 If general or group factors are found, what is their psychological nature? In other words, the problem becomes one of naming and describing in psychological terms the ability or abilities, character, or characters of performance intelligence.

7 What are the implications of the findings of this study for the theory of intelligence and for the theory and technique of performance testing?

C Method

A group of 34 of the most commonly used performance tests were administered by the author to 56 nine-year-old boys in a New York City grade school These subjects were selected to include the entire male population of this school within a three month range of nine years, six months, that is, from 9 years, 3 months to 9 years, 9 months. This particular school draws principally from the Hebrew population and represents the middle class economic level The largest percentage of the subjects (59%) were selected from Grade *4B*, with 16 per cent each from Grades *4A* and *5A* and the smallest number (9%) from *3B*. As measured by the Henmon-Nelson Intelligence Test, the mean *IQ* was 112, standard deviation of 16 The range of these *IQ*s was from 80 to 142 The mean score on the Brown Personality Test was 20.96 which, according to the norms supplied by the author (2) of the test, places the average of the group in the sixth decile, tenth indicating poorest adjustment. The median of our population, 18 5, is slightly below the median (17 1) of 2748 unselected cases, aged 9 to 14, upon whom the test was standardized. The scores of the selected subjects which range from 3 to 48 cover the entire standardized range from "excellent adjustment" to "very poor adjustment." In mechanical aptitude as measured by the Revised Minnesota Paper Form Board a median of 21.37 places this group in the 60th percentile as reported by the authors of the test for a group of 108 nine year old boys (10).

The above, then, indicates that we are dealing with a highly homogeneous group both in age and school grade with sex held constant, and above average in intelligence and mechanical ability in which we are primarily interested, but slightly below normal in personality adjustment. Since it has been pointed out that there was a preponderance of Hebrew children and since this is a city group exclusively this may merely mean that the norms for the personality tests are not adequate It must be remembered also that the intellectual and mechanical ability norms hold age constant while the measure of adjustment is based on a wide and uncontrolled age range.

The tests selected included the entire Pintner-Paterson Scale of Performance Tests, the Maxfield Color Cubes Test, Lincoln Hollow Square, Witmer Form Board, Witmer Cylinder Test, Dearborn Form Board 3, Porteus Maze, Revised Minnesota Paper Form Board Test, Brown Personality Test, and Henmon-Nelson Intelligence Test In addition, a number of other tests were administered which had been selected or devised by the author to measure certain assumed mental traits They were· Aiming, Speed of Reaction, Eye-Hand Coördination Maze, Form Cancellation, Dot Counting, Circle Block Form Board, Square Block Form Board, Form Card Sorting, Manual Dexterity, and Visual Acuity.

The Aiming test is similar to Whipple's save that it is one-half the size and the subject is instructed to sit with the test on his desk and thrust from eye level downward, rather than standing and thrusting forward from shoulder level It was felt that the former procedure more nearly duplicated the type of aiming behavior to be found in the performance test situation. The Speed of Reaction test was simply the Two Plate Tapping test The Dot Counting and Form Cancellation tests are described in Whipple's Manual (13)

The Square and Circle Block Form Boards were home made and were designed to measure speed of reaction as applied to Form Board performance Each consisted of a board the size of the Seguin board with recesses and blocks of the same size as the Seguin, save that the Circle Block test consisted of 12 circular recesses and circular blocks, while the Square Block test was composed of 12 square recesses and Blocks The Form Card Sorting test was a card sorting apparatus of 15 compartments The 15 test cards each bore some design commonly used in form boards, square, circle, diamond, semi-circle, triangle, etc The test series consisted of two trials with 45 cards to be sorted in each trial The Manual Dexterity test is the Finger Dexterity test devised by Johnson O'Connor with instructions adapted by Whitman (14) for use with children. The Visual Acuity test is the well-known Snellen chart. The Eye-Hand Coordination maze was a simple paper-and-pencil square maze The alley was of equal width about the four sides of the square, but decreased a constant amount as the centre was approached Thus, a measure of coordination could be obtained in four directions

These tests were administered at three different testing periods at no greater than two-week intervals The data thus collected, yielding some 75 variables, were coded and punched on Hollerith cards The medians, means, and standard deviations of the uncoded distributions were computed and compared with those reported in the literature This comparison revealed that the distributions of the present study are statistically similar to those obtained by other experimenters Where discrepancies are apparent, there are as wide or wider differences between other experimenters as between our results and those of others, indicating inherent unreliabilities of the measures or of the samples The intercorrelations were obtained for all the variables, but of this number only 33 were selected for the factor analysis Only those variables were chosen which satisfied any two of the following criteria

1 The variable must discriminate individual differences Of several variables obtained from administering the same test, that one was selected which yielded the largest standard deviation

2 The variable must be free from obvious contamination by other factors. For example the score on the Substitution Test is contaminated by weighting the time record by the subjective estimate of errors

INTERCORRELATIONS OF THE THIRTY-THREE VARIABLES SELECTED FOR THE CENTROID ANALYSIS. N=56

Variable Number	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10	-11	12	-13	14	15	-16	-17	-18	-19	-20	21	22	-23	-24	-25	-26	-27	-28	-29	-30	31	-32	33
-1		39	54	35	50	12	10	31	31	45	37	09	20	-11	-13	38	42	12	23	11	05	20	-02	04	03	-03	07	13	33	44	19	09	25
-2	39		42	24	32	21	12	13	14	36	35	12	34	11	02	27	45	20	32	40	29	11	18	34	12	10	-03	13	15	21	18	06	27
-3	54	42		33	51	22	26	23	19	17	16	16	24	-04	-04	20	49	27	33	42	17	13	07	-03	05	-02	01	07	10	34	15	13	37
-4	35	24	33		22	21	07	06	11	16	33	07	-02	16	07	25	34	14	06	22	14	20	04	-04	-15	-15	-10	-07	17	02	-02	06	06
-5	50	32	51	22		00	07	04	27	42	17	15	24	-16	09	16	42	21	35	45	11	05	-02	19	06	01	09	07	23	33	24	-01	37
-6	12	21	22	21	00		08	-27	09	27	06	-01	-08	-01	-34	01	-02	-09	-05	04	10	32	06	-03	12	-12	-12	01	-04	-11	-11	07	06
-7	10	12	26	07	07	08		22	-01	-01	11	-09	06	02	08	06	08	-07	14	14	-00	02	01	-06	-10	-14	-03	06	02	-14	-12	00	00
-8	31	13	23	06	04	-27	22		-22	-08	03	-29	01	-06	-02	28	18	11	18	01	-14	-09	-00	-12	-15	-14	-22	06	-00	-04	14	-09	-05
-9	31	14	19	11	27	09	-01	-22		37	15	12	25	03	-06	13	-03	19	14	11	13	18	22	-00	07	22	-11	14	26	33	26	13	06
-10	45	36	17	16	42	27	-01	-08	37		21	13	16	-22	-09	06	30	10	01	28	19	18	-09	01	-13	04	-04	13	25	16	14	-16	31
-11	37	35	16	33	17	06	11	03	15	21		12	40	19	25	36	38	-00	-12	26	13	-01	14	18	17	06	26	02	40	22	20	-06	24
12	09	12	16	07	15	-01	-09	-29	12	13	12		28	06	22	17	20	04	-06	33	23	01	16	33	-11	25	46	05	14	28	-12	-05	33
-13	20	34	24	-02	24	-08	06	01	25	16	40	28		12	14	19	25	20	03	40	33	05	18	32	30	16	27	18	29	46	43	02	13
14	-11	11	-04	16	-16	-01	02	-06	03	-22	19	06	12		15	12	09	-10	-23	-09	07	04	02	-04	-09	-04	12	-00	-01	04	16	15	-14
15	-13	02	-04	07	09	-34	08	-02	-06	-09	25	22	14	15		16	18	-21	-10	24	28	-28	15	21	-18	16	22	-00	09	11	-10	-00	24
-16	38	27	20	25	16	01	06	28	13	06	36	17	19	12	16		37	-01	-02	40	05	-12	34	13	-11	-08	27	18	28	18	07	24	33
-17	42	45	49	34	42	-02	08	18	-03	30	38	20	25	09	18	37		14	00	40	19	06	12	19	-02	-04	16	09	27	32	12	09	30
-18	12	26	27	14	21	-09	-07	11	19	10	-00	04	20	-10	-21	-01	14		12	19	05	07	27	-02	11	-11	-06	-11	07	16	18	-26	05
-19	23	32	33	06	35	-05	14	18	14	01	-12	-06	03	-23	-10	-02	00	12		24	12	14	-07	30	09	15	-16	10	01	19	05	09	08
-20	11	40	42	22	45	04	14	01	11	28	26	33	40	-09	24	40	40	19	24		31	11	23	34	00	20	30	15	40	18	17	-08	45
21	05	25	17	14	11	10	-00	-14	13	19	13	23	33	07	28	05	19	05	12	31		32	27	22	-04	-01	09	01	08	24	-03	-22	15
22	20	11	13	20	05	32	02	-09	18	18	-01	01	05	04	-28	-12	06	07	14	11	32		-00	39	12	18	-07	24	13	33	12	-16	08
-23	-02	18	07	04	-02	06	01	-00	22	-09	14	16	18	02	15	34	12	27	-07	23	27	-00		22	20	18	17	20	29	22	-05	06	32
-24	04	34	-03	-04	19	-03	-06	-12	-00	01	18	33	32	-04	21	13	19	-02	30	34	22	39	22		37	42	30	35	22	38	01	-11	30
-25	03	12	03	-15	06	12	-10	-15	07	-13	17	-11	30	-09	-18	-11	-02	11	09	00	-04	12	20	37		20	13	11	19	38	24	-01	06
-26	-03	10	-02	-15	01	-12	-14	-14	22	04	06	25	16	-04	16	-08	-04	-11	15	20	-01	18	18	42	20		-01	13	28	38	11	06	11
-27	07	-03	01	-10	09	-12	-03	-22	-11	-04	26	46	27	12	22	27	16	-06	-16	30	09	-07	17	30	13	-01		20	24	34	01	00	44
-28	19	13	07	-07	07	01	06	06	14	13	02	05	18	-00	-00	18	09	-11	10	15	01	24	20	35	11	13	20		35	23	17	-09	23
-29	33	15	10	17	23	-04	02	-00	26	25	40	14	29	-01	09	28	27	07	01	40	08	13	29	22	19	28	24	35		35	28	10	20
-30	44	21	34	02	33	-11	-14	-04	33	16	22	28	46	04	11	18	32	16	19	18	24	33	22	38	38	38	34	23	35		26	14	34
31	19	18	15	-02	24	-11	-12	14	26	14	20	-12	43	16	-10	07	12	18	05	17	-03	12	-05	01	24	11	01	17	28	26		15	-04
-32	09	06	13	06	-01	07	00	-09	13	-16	-06	-05	02	15	-00	24	09	-26	09	-08	-22	-16	06	-11	-01	06	00	-09	10	14	15		02
33	29	27	37	06	37	06	00	-05	06	31	24	33	13	-14	24	33	30	05	08	45	15	08	32	30	06	11	44	23	20	34	-04	02	

Note: For an identification of the variables, see Page 19, Table II. The PEs range from .06 to .09, as shown by the Table below.*
; The intercorrelations of the sub-tests of the Pintner-Paterson Performance Scale are those included within the boxed portion of the Table, i.e., variables 1 - 15.

*PE	.09	.08	.07	.06
r	.00 - .23	.24 - .40	.41 - .52	.53 - .62

TABLE 1
CORRELATIONAL MATRIX

3 Whenever possible, that variable was selected which was free from subjectivity in obtaining it Whenever time and errors, or time and moves were recorded for a test the less subjective score, time, was taken

The variables selected are listed in Table 2

The correlational matrix shown in Table 1 was factored by Thurstone's centroid method (12) into the factorial matrix of four factors The loadings and communalities of these factors are shown in Table 2 However, before commencing the analysis, the signs of all variables in which negative correlations would be expected (i e., time, errors, moves, and the like) were reflected The necessity of this sign changing could have been remedied if such variables had been distributed at the outset in increasing order with the lowest score indicating best performance, but it was felt that the procedure

TABLE 2
Factorial Matrix

DESCRIPTION OF THE VARIABLES WITH THEIR RESPECTIVE FACTOR LOADINGS

Variable Number	DESCRIPTION OF VARIABLE	FACTOR LOADINGS I	II	III	IV	Communality $h^2_{k\,1234}$
1	Mare and Foal Time	592	-491	289	257	763
2	Seguin Form Board Time of shortest trial	620	-286	-040	-184	502
3	Five Figure Form Board Time	591	-460	-026	-166	589
4	Two Figure Board Time	321	-389	-167	201	323
5	Casuist Form Board Time	556	-259	187	-104	422
6	Triangle Test Time	082	-197	-092	196	094
7	Diagonal Test Time	100	-252	-135	-156	116
8	Healy Puzzle A Time	028	-400	123	-220	224
9	Manikin Test Time	377	014	258	212	254
10	Feature Profile Time	375	-163	376	181	234
11	Ship Test Time	508	-023	-221	179	337
12	Picture Completion Score	351	301	-255	019	301
13	Substitution Test Time	580	243	-004	-065	400
14	Adaptation Board Score	046	088	-314	289	192
15	Cube Test Number Correct	154	248	-440	-221	328
16	Warfield Color Cubes Design C Time	473	-144	-305	170	366
17	Witmer Form Board Shortest Trial Time	580	-295	-196	031	463
18	Lincoln Hollow Square Total Time	186	-076	162	-206	109
19	Witmer Cylinder Shortest Trial Time	251	-175	337	-429	391
20	Dearborn Form Board Total Moves	645	017	-190	-288	535
21	Porteus Maze Years Credit (M.A.)	353	177	-205	-167	226
22	Speed of Reaction dominant hand number of taps	273	100	304	154	201
23	Aiming dominant hand total error	366	236	-163	-145	237
24	Eye-Hand Coordination Plate Time	464	423	090	-241	460
25	Form Cancellation Time	198	256	267	-079	182
26	Dot Counting Total Time	244	449	253	-070	330
27	Dot Counting Total Errors	305	404	-328	136	382
28	Circle Block Form Board Shortest Trial Time	322	243	188	081	205
29	Square Block Form Board dominant hand Time	543	161	126	227	368
30	Form Card Sorting Total Time	644	298	271	089	505
31	Manual Dexterity Pegboard A dominant hand	321	027	314	133	220
32	Visual Acuity	053	-066	-117	221	070
33	Revised Minnesota Paper Form Board Total Score	576	053	-165	-092	326

of distributing all variables in decreasing order would cause less confusion both in computation and in interpretation The numbers of all variables in Table 1 which were reflected are preceded by a negative sign, and hence should be thought of as negative variables

The factors shown in Table 2 are not psychologically meaningful. In order to obtain a meaningful solution any one of several methods must be employed for rotating the reference axes. (The factor loadings originally obtained are merely the projections of the test vectors upon these axes) The method described by Guilford (5) was chosen. This is a graphical method whereby two factors are dealt with at a time, "rotating them about a third axis as a pivot, computing new factor loadings after each rotation, and repeating the process as many times as is necessary," that is, until further rotations would not decrease the number of negative loadings

TABLE 3

ROTATED LOADINGS OF THE FIRST THREE FACTORS WITH THEIR COMMUNALITIES, AND AUGMENTED LOADINGS

Variable Number	ROTATED LOADINGS I	II	III	Communality h^2_{k123}	AUGMENTED LOADINGS I*	II*	III*	Communality h^2_{k123}
1	-089	726	374	675	-108	884	455	1.0001
2	098	422	529	460	143	617	774	1.0002
3	-064	480	572	561	-085	610	763	.9550
4	-151	198	469	282	-284	373	883	.9595
5	085	545	326	411	133	851	509	1.0010
6	-120	049	192	054	-518	211	829	1.0001
7	-157	047	254	092	-519	155	840	.9990
8	-320	253	097	176	-763	603	231	.9991
9	217	398	056	209	475	871	123	.9994
10	051	400	196	202	114	892	437	.9996
11	257	129	473	307	464	233	854	.9989
12	444	-132	294	301	809	-241	536	.9996
13	520	234	266	355	826	372	423	.9996
14	099	-239	203	108	301	-727	617	.9994
15	292	-331	290	279	553	-627	549	1.0009
16	137	091	557	337	236	157	959	1.0000
17	068	287	612	462	100	422	901	.9997
18	037	252	038	067	144	979	147	1.0008
19	-010	455	002	207	-022	1.000	004	1.0005
20	365	215	522	452	543	320	776	.9994
21	341	-018	286	198	766	-040	642	1.0005
22	232	342	-074	177	553	814	-176	.9994
23	397	-002	241	216	655	-004	519	1.0004
24	607	173	058	402	958	273	091	1.0006
25	322	216	-158	176	769	516	-377	.9997
26	509	161	-159	325	893	283	-349	.9993
27	505	-220	246	364	837	-365	408	.9993
28	379	232	-023	190	852	521	-032	1.0000
29	431	340	189	337	742	586	325	.9996
30	601	454	059	577	791	598	130	1.0002
31	197	403	-021	202	433	897	-047	.9995
32	-026	-033	138	021	-180	-229	957	1.0007
33	336	159	423	317	597	282	751	.9999

Although the analysis was carried to four factors, the fourth factor was ignored when the axes were rotated, since the fourth factor loadings appeared insignificant according to Guilford's criteria Only one loading, that for variable 19, contributes more than 100 to the communality

In order to express graphically the relationships among these three factors and thus clarify an interpretation of them, their loadings were augmented by multiplying the rotated loadings by the reciprocal of the square root of the communality The communalities will then approximate unity as seen in the last column of Table 3 The augmented factor loadings may now be used in plotting the tests as points on the surface of a sphere of a radius equal to unity This was done by plotting the augmented factor loadings III against II, as shown in Figure 1 The first factor is then perpendicular to the plane of the diagram The spherical triangle *ABC* has been drawn to show the location of the points in three planes

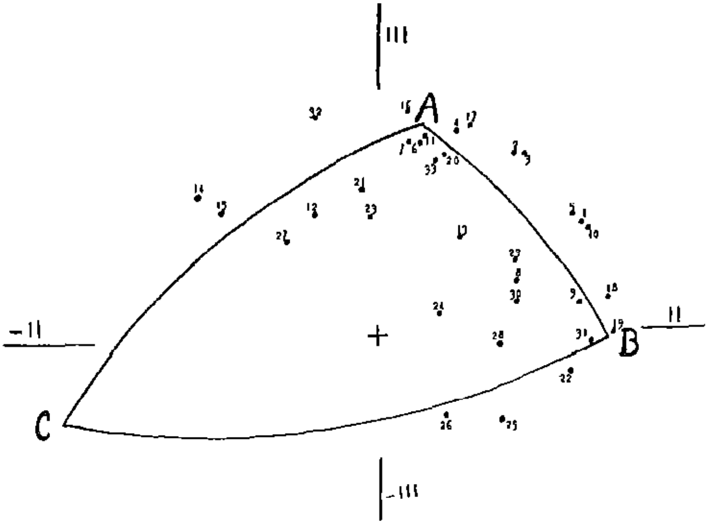


FIGURE 1
AUGMENTED LOADINGS PLOTTED AS POINTS ON A SPHERE

D RESULTS

1 *Interrelationships*

The interrelationships found to obtain among performance tests of intelligence is seen from an examination of the correlation coefficients presented in Table 1 When due account is taken of the nature of the score correlated, and the table is so adjusted as to make all coefficients alike as regards their expected sign (by reflecting signs as previously described), negative correlations are still found among the tests. This is contrary to the expectation of positive correlations as hypothecated earlier, and is particularly important in view of the fact that all of these tests purport to measure intelligence or some characteristic or attribute of intelligence For instance a glance at the coefficients of the tests in the Pintner-Paterson Scale indicates that the Adaptation Board and the Casuist Board are negatively related. The same statement seems to hold for two such apparently similar tests as the Triangle Test and the Healy Puzzle *A*. Similar conditions exist between the Triangle Test and the Knox Cubes, the Healy Puzzle *A* and the Manikin, the Healy Puzzle and the Picture Completion, and the Feature Profile and Adaptation Board. While other negative coefficients are found among the tests of the Pintner-Paterson Scale they are so low as to indicate not antithesis but rather no relationship

Positive coefficients of moderate magnitude do obtain among many of the tests as for example, Mare and Foal and Five Figure Board, Mare and Foal and Casuist, Five Figure Board and Casuist, and many others. However, since these same generalizations will appear from later data (the Factorial Analysis) it seems unnecessary to go further into any summary of the table of intercorrelations Suffice it to say that the relationships range from moderately high positive through no relationship to moderately high inverse relationship, and that it is difficult to predict from the names and descriptions of any two tests just which relationships may be found between them

2. *Factorial Analysis*

An examination of Tables 2 and 3 reveals the answers to several of the questions raised earlier and sheds some light on others First, the existence of a general factor seems not to be indicated The assumptions and mathematics of the Centroid Factorial Analysis are such that to support the claim of a general factor all loadings for Factor I must be positive and moderately high. For a factor to contribute as much as .100 to the communality the loading must be 316. An inspection of the Factor I loadings in Table 2 indicates that not only are many of the positive loadings less than .316 but many of them are actually negative.

By inspecting each test in Table 3, Rotated Loadings, and noting with which factor each has the highest loading, it is possible to observe certain groupings of tests The Picture Completion, Substitution, Cube, and Porteus Maze all have higher positive loadings for the first factor than for either of the other two However the number of tests in which Factor I does not play a part (as indicated by practically zero or even minus loadings) is also large An inspection of the loadings of Factors II and III indicates at least two more group factors running through certain of the tests For instance the presence of a group factor (II) is indicated in the grouping of Mare and Foal, Casuist, Healy, Manikin, Feature Profile, Lincoln Hollow Square, and Witmer Cylinder The tests having high loadings with Factor III may be grouped as follows Seguin, Five Figure, Two Figure, Triangle, Diagonal, Ship, Adaptation Board, Maxfield Cubes, Witmer Form Board, Dearborn, Minnesota Paper Form Board

Beyond three factors all that can be obtained are factors having a significant loading for only one test, as indicated by the fact that Factor IV has but one significant loading (–429), that for variable 19, the Witmer Cylinder test; the others are all less than 316, and therefore, as stated above, do not make a significant contribution to the total communality Hence, factors beyond the third are not indicative of general or group characteristics but rather of uniquenesses in each test.

There seem to be no tests which are unique, that is, do not have relatively high positive loadings for at least one of the three factors. That all of the tests have some unique character in addition to their group character or characters is indicated by the fact that none of the communalities computed from unaugmented loadings for three factors approach unity. However the difference between obtained communalities and unity may be a function of the unreliability of the tests.

Although the triangle *ABC* (Figure 1) was not drawn with any model or category of triangles in mind, the best fit very closely approximates a right spherical triangle. This is a primary assumption of factor analysis, that mental traits are orthogonal, or in other words, are primary independent abilities

A first glance at the figure reveals that simple structure does not obtain, i.e., the tests do not fall exactly on any angle of the triangle. This confirms the statement that no test is entirely unique. All the tests do not cluster in definite groups nor do they distribute themselves along the planes About half of the tests seem to cluster into two primary groups at *A* and *B*. The remainder apparently represent measures of more complex abilities. It is interesting to note that a number of the tests which were selected or devised to measure certain so-called pure traits, such as the Eye Hand Coordination Maze (No 24) and Aiming (No. 23) are among this group of complex measures

First, let us consider the cluster at *A* Here are found four form boards from the Pintner-Paterson series (Nos 4, 6, 7) and the Ship Test (No. 11) which although logically a picture completion test, is reacted to by the subjects of the present experiment, at least, as a form board The first reaction seems to be to fill the large recess with the 10 equal blocks upon which appear the cut-outs of the ship picture Often only after he has filled the recess does the subject comprehend the total problem of re-arranging the blocks to form a picture of a ship Also found grouped at *A* are the Witmer Form Board (No. 17), the Dearborn Form Board (No. 20), the revised Minnesota Paper Form Board (No 33), and the Mayfield Color Cubes Design *C* (No 16) These tests all appear to involve an ability that requires the perception of form and space relations, a trait which Thurstone has identified as Visualizing This confirms our inference of group Factor III

At *B* we find a small cluster of tests, (Nos 9, 18, 19, 22, 31). Test 22 is the speed of reaction or tapping test and is probably as pure a measure of speed as it is possible to obtain The Manual Dexterity Test (No. 31) which is also found grouped here is a controlled speed test in that the subject is required to react quickly in a regulated fashion Test 9 (the Manikin Test) which is too simple for this group of subjects, becomes a speed test The Lincoln Hollow Square (No 18) is heavily weighted by the speed factor in its less difficult parts It will be remembered that this test is composed of 10 problems of increasing difficulty The first six or seven parts are very easy to complete and are thus measuring speed of reaction to a large extent. The remaining parts were also easy for most of our subjects The score on the Witmer Cylinder Test

(No 19) is the time of the shortest trial and it becomes consequently a matter of finding the desired cylinder quickly rather than of perceiving form and space relationships as each cylinder is reacted to These tests all seem to involve an ability to make a quick identification of simple perceptual material when mixed with like or similar material, and for this reason may be thought of as Perceptual Speed (Thurstone) This description gives psychological meaning to our group Factor II.

Along the plane *AB* are found tests more complex than those at *A* or *B* and dependent upon them A consideration of the makeup of these tests shows the logic of this statement The Seguin Board (No 2) and Five Figure Board (No. 3) were easier for our subjects than the other tests grouped at *A*, and therefore would seem to involve more of the speed factor Tests 1, 5, 10 (Mare and Foal, Casuist Board, and Feature Profile) lie relatively close to the speed cluster at *B*, but at the same time there was inherent in their makeup a more definite form of relationship to be perceived than is true of the other tests at *B*

There is no grouping of tests at *C*, indicating that none of the tests selected for the present analysis is heavily weighted with this trait However, when we inspect the figure further and note the tests lying in the planes originating from *C* we find a suggestion as to the identification of this factor. The Dot Counting Test, of which we have two variables included; time (No 26) and errors (No 27) lie near two different *C* planes. The Time measure lies in the *CB* plane or in other words involves speed of reaction to an extent, while the accuracy measure (errors) lies in the *CA* or form perception plane The Dot Counting test is a test where the subject counts a linear or clustered grouping of dots The linear groupings may be grouped in threes or fours and the like Superior performance is gauged by the subject's ability to accommodate himself to the objective grouping of the cards Thus, excellence of this performance depends upon the ability to discover and make use of this principle Tests 14 and 15 (Adaptation Board and Cube Imitation) call for this same behavior but are also dependent upon space perception as seen by their proximity to the *CA* plane Tests 12, 21 and 23 (Picture Completion, Porteus Maze, and Aiming) may be described in similar terms as being dependent upon the ability to discover a pattern or rule of procedure. The Porteus Maze is usually described as a measure of planning ability but it is shown here to be measuring a complex trait, which may be planning ability, but which also involves at least two simpler abilities The Aiming test was selected for our battery because it was thought that it measured aiming However, when we control the subject's aiming by restricting him to aim in a sequence, or as it developed, a pattern, we are also obtaining a measure of the instructions to follow this pattern or rule We would, therefore, name this trait, and group Factor I which it describes, Induction (Thurstone).

Further confirmation for naming the points of our triangle in the above terms is obtained from noting the location of the Circle and Square Block Tests (Nos. 28 and 29) They are both speed tests and are found clustered near *B* but since the placing of a square block in a square recess requires a greater degree of form discrimination than the placing of a circle block in a circular recess, we would expect to find the Square Board to be nearer the Visualizing plane than Induction. Such is the case.

Finally, it is our purpose to show that since the attempts to measure intelligence as an entity have not yielded a working definition, the way is now open through factor analysis to redefine our conception of intelligence in terms of interrelations of measurable unitary abilities The present investigation has set out to discover those factors which operate in differentiating individual performances in a group of non-verbal tests of mental ability

The evidence was presented above that three is the smallest number of independent abilities that need be postulated to account for the table of intercorrelations of a group of performance tests Further experimentation with these and other tests will probably increase this number

There is a very plausible explanation for the absence of a general factor. There is no clear-cut general factor because all of the tests are not actually or theoretically measuring a general ability Certain tests were purposely selected because they were thought to measure a very specific ability, for example, speed of movement, and speed of reaction. The fact that there is an indication of induction as a factor suggests that here we are measuring an ability, the ability to "discover some principle or rule that governs the material," which is very similar to Spearman's general intellective factor A test devised by Spearman to measure "*g*" was one of the tests used by Thurstone and found by him to involve this trait which he has named Induction. Bernreuter, under whose direction Murphy's study was pursued, states that her factor "labeled the 'mental manipulation of relationships expressed symbolically' probably corresponds with Thurstone's 'induction'" (1).

There is also a similarity between our other two factors and those reported by Murphy, which she found to operate in paper-and-pencil performance and intelligence tests. The similarity of Thurstone's Visualizing and Perceptual Speed factors to Murphy's "Mental Manipulation of spatial relations" and "Speed of Eye-Hand Coordination" has also been pointed out by Bernreuter.

These evidences of the similarity of the psychological natures of the differentials of intelligent behavior as found by three different investigators, utilizing two approaches to measurement, i e, performance and paper-and-pencil tests, compel us to redefine our terms for the description of mental ability, whether it is the ability to think in abstract terms or to adjust to one's environment, or whether it is the capacity to acquire or the general ability consciously to adjust thinking to meet new requirements. We are now in a position experimentally to determine whether it is any or all of these things working together, or whether it is an entirely different interaction of abilities

The findings of the present study give us a foundation on which to build new tests for the measurement of the primary independent characteristics of intelligence We will need measures, each of which will sample one trait or factor, and with the inclusion of these in a new battery of performance tests we will have tests representative of a sampling of performance abilities

E SUMMARY AND CONCLUSIONS

To further the study of factors which may be assumed to account for Performance Test behavior, a battery of well-known tests were administered to a group of 56 nine-year-old boys in a New York City grade school The ages of these subjects ranged within three months of nine years six months (sigma = 2 months). The tests selected included the Pintner-Paterson Battery, the Seguin-Witmer-Sylvester Form Board, the Lincoln Hollow Square, the Witmer Cylinder Test, the Dearborn Form Board 3, the Porteus Maze, Revised Minnesota Paper Form Board Test, Henmon-Nelson Intelligence Test, the Brown Personality Test and several other tests devised by the author or selected to measure assumed psychologically primary, independent abilities.

The data collected from scoring these tests yielded 77 variables. The means, medians, and standard deviations were computed for each distribution to indicate comparability of our data to standard usage of the tests. Intercorrelations for all 77 variables were computed and 33 variables were selected for a factorial analysis This correlational matrix was factored by Thurstone's Centroid Method into the factorial matrix of three factors. In order to obtain a factorial solution, psychologically meaningful, the reference axes were rotated, the new loadings augmented and plotted as points on the surface of a sphere.

The following conclusions may be drawn from an analysis of these data:

1. The interrelationships among the performance tests used were found to range from moderately high positive through no relationship to moderately high inverse relationship.

2. The presence of a general factor in the correlational matrix was not disclosed

3. Several group factors or abilities were found to account for the common elements in performance as indicated by the correlational matrix.

4 It seems evident from the findings that the tests used in this battery are largely measures of two primary abilities with some evidence of a third ability, not well measured by any single test.

5. There seem to be no tests which do not have some unique character in addition to their group character in view of the fact that none of the communalities approach unity.

6. The psychological nature of the three group factors found may be described as (*a*) an ability that requires the perception of form and space relations, (*b*) the ability to make a quick identification of simple perceptual material when mixed with like or similar material, and (*c*) the ability to discover a pattern or rule of procedure. These traits have been previously described in similar terms by Thurstone and identified by him as Visualizing, Perceptual Speed, and Induction, respectively.

7. The mental traits as measured by the test battery of the present investigation are orthogonal, or in other words, they are independent mental abilities.

REFERENCES

1 BERNREUTER, R. G (Editor) *Clinical Psychology*

2 BROWN, F Suggestions with Regard to Use and Interpretation of the Brown Personality Inventory Privately mimeographed, Alfred Willson's Children's Center, Columbus, Ohio

3 GAW, F A study of performance tests *Brit J Psychol*, 1924, **15**, 374-392

4 GODDARD, H H The adaptation board as a measure of intelligence *Train Sch Bull*, 1915, **11**, 182-188

5 GUILFORD, J P Psychometric Methods New York, McGraw-Hill, 1936

6 HUTT, R B W Standardization of a color cube test *Psychol Clin*, 1925, **16**, 77-97

7 MURPHY, LAURA W The relation between mechanical ability tests and verbal and non-verbal intelligence tests *J of Psychol*, 1936, **2**, 353-366

8 PASCHAL, F C The Witmer Cylinder Test Hershey, Pa Hershey Press, 1918

9 PINTNER, R, & ANDERSON, MARGARET M The Picture Completion Test Baltimore Warwick and York, 1917

10 QUASHA, W H, & LIKERT, R The revised paper form board test *J Educ Psychol*, 1937, **28**, 197-204

11. THURSTONE, L L A new conception of intelligence *Educ Rec*, 1936, **17**, 441-450.

12. ——— The Vectors of Mind Chicago Univ of Chicago Press, 1935

13 WHIPPLE, G M Manual of Mental and Physical Tests, Volume I, Simpler Processes. Baltimore Warwick and York, 1924

14 WHITMAN, ESTHER C A brief test for manual dexterity *J Educ Psychol*, 1925, **16**, 118-123

15 WOOLLEY, H T., & FISCHER, C R Mental and physical measurements of working children *Psychol Monog*, 1914, **18**, No 77 Pp 247

16 WORTHINGTON, MYRTLE R A study of some commonly used performance tests *J Appl Psychol*, 1926, **10**, 216-227

The Woods Schools
Langhorne, Pennsylvania

The Journal of Genetic Psychology, 1939, 54, 107-118

RELATIVISM IN THE THINKING OF SUBNORMAL SUBJECTS AS MEASURED BY CERTAIN OF PIAGET'S TESTS*[1]

Yale University and Letchworth Village

ELIZABETH B LANE AND ELAINE F. KINDER[2]

A THE GENERAL PROBLEM

The chief principle employed by Piaget to explain the characteristics of children's thinking is that of egocentricity Egocentricity is Piaget's name for the fact that the child's thinking cannot go beyond his own immediate point of view. The child does not question his own impression: the world is simply what it appears to him to be, because he does not yet know that portion of it which is independent of him. He does not understand that the world may appear in a different light to others. Piaget contends that it is "because he has always taken his own point of view as something absolute that the child remains ignorant of the habits of relativity and comparison" (3), and accordingly reduces ideas of relationship to ideas of simple quality. This may be illustrated by Piaget's further statement that the child "does not realize that a brother must necessarily be the brother of somebody, that an object must necessarily be to the right or left of something, or that a part must necessarily be part of a whole, but thinks of all these notions as existing in themselves, absolutely" (3, p 131). The increase in ability to grasp ideas of relationship is termed by Piaget the growth of relativism Relativism grows as egocentricity declines and as the child becomes increasingly able to abandon his own immediate point of view. These processes are conceived as being continuous in the developing child, but Piaget believes that certain stages or levels stand out and can be described. These levels he has defined in terms of chronological

---

*Received in the Editorial Office on April 2, 1938

[1] Paper read before the Eighth Annual Meeting of the Association of Consulting Psychologists, Albany, New York, May 8, 1937.

[2] The writers express appreciation to Dr James Q Holsopple for his reading of the manuscript and suggestions

age, and he has emphasized the rôle of social intercourse in effecting the transition from one level to the next higher

Piaget's work has been criticized as being based on experimental evidence which is too limited for his far-reaching interpretations.[3] Attention has been called to the small numbers of subjects used in his studies, and to the need for applying his methods to other groups of similar make-up. Also, the importance of extending these techniques to work with non-typical individuals, such as mental defectives or psychotics, should not be overlooked.

The use of Piaget's methods with subnormal subjects is of special significance because in mental defectives we have a group of individuals for whom the increase in chronological age and the growth of "mental maturity" sustain a relationship which differs from that for the majority of persons. This fact has considerable significance for the problems studied by Piaget Piaget has described the thinking of the "eight-year-old child," for example, but whether this child thinks as he does because he has lived eight years, or because he has attained a level of mental maturity usually but not necessarily reached by individuals who have lived eight years, remains a question. Assuming that Piaget has demonstrated that the ability to handle certain ideas increases for many children as they grow older, this fact alone does not lead, on either theoretical or empirical ground, to the conclusion that chronological age is the sole or chief determinant of this process. If satisfactory methods of studying them can be established, factors of intellectual or social maturity, to a certain extent independent of chronological age, may be found to have an important or even major influence. The examination of subjects who are mentally defective offers a possible means of comparing the two factors of intellectual maturity and chronological age with respect to their relative influence on the growth of the abilities investigated by Piaget. The present study was undertaken as an initial step in work with this problem

It is clear that in order to attack this problem the factors of chronological age and intellectual maturity must be controlled and varied independently; and it is also clear that some measure of intellectual maturity must be used to do this. The term "mental age" has until now been avoided, deliberately. It is the authors' opinion that the uncritical adoption of the familiar contrast of "mental" with "chronological" age tends to contribute to the practice of accepting as final the results of certain techniques for measuring intelligence It is difficult, at best, to avoid attributing to particular units for measuring mental maturity an objectivity comparable to that of the units by which we measure chronological age

---

[3] Those who make this criticism seem sometimes not to do justice to the general significance of Piaget's contribution to genetic psychology and to the psychology of the thought-processes. It may be pointed out that any experimentation dealing with particular aspects of Piaget's work is in a sense a tribute to the originality and insight of his thinking

There is another reason for minimizing the use of the term "mental age" here. Let us suppose that a relationship should be found between the development of the abilities studied by Piaget and the development of abilities used in measuring intelligence If these latter abilities are regarded as constituting "mental age" in some special and ultimate sense it becomes dangerously easy to conclude that the development of these "mental age" abilities causally determines the abilities studied by Piaget. It is our contention, however, that our present knowledge does not warrant this conclusion. There is no reason for asserting that the abilities now used in testing intelligence are more fundamental to intelligence than others which might be found to be related to them It is entirely possible, for example, that the measurement of the amount of relativism in a subject's thinking should have a practical value for predicting his mental level, and a theoretical value for understanding what that mental level means, that would render it of as much importance as the measurement of the more familiar abilities long associated with intelligence tests.

The above discussion is not intended to depreciate the practical or theoretical value of the concept of "mental age" or of any instruments for measuring intelligence. Throughout the present experiment the results of the Stanford-Binet Scale, being readily available, were used as our determinants of the mental maturity of our subjects. These results are, of course, expressed in terms of mental age, and the term will be used frequently from now on in reporting our procedure and results, but its use should be considered in the light of the foregoing qualifications.

B Procedure

The present study attempts to reproduce with subnormal subjects one of Piaget's experimental techniques, selected because of its interest for the problem of egocentricity and relativism, and because of its simple and definite nature. This is a series of 12 questions dealing with ideas of relation, six with the concepts of "brother and sister," and six with those of "right and left." These questions are listed below.

TESTS IN "THE LOGIC OF RELATIONS"*

I *Brothers and Sisters*

1 How many brothers have you? And how many sisters? (If the subject has a brother *A* and a sister *B*)—how many brothers has *A?* And how many sisters? How many brothers has *B?* And how many sisters?

2. How many brothers are there in your family? How many sisters? How many brothers and sisters altogether?

3. There are three brothers in a family John, Dick, and Tom How many brothers has John? And Dick? And Tom?

4 Are you a brother (or a sister)? What is a brother (or sister, according to the sex of the subject)?

5 George has three brothers, Paul, Henry, and Charles How many brothers has Paul? And Henry? And Charles?

6 How many brothers are there in this family?

II. *Right and Left*

7 Show me your right hand, your left Show me your right ear, your left.

8 Show me my right hand, my left Show me my right ear, my left.

9. (A penny and a pencil are placed before the subject) Is the pencil to the right or to the left? And the penny?

10. (Examiner holds a penny in her right hand and has a watch on her left arm) Have I got the penny in my right hand or in my left? And the watch?

11 (A pencil, key, and penny are placed before the subject.) Is the pencil to the left or to the right of the key? And of the penny? Same for other objects

12. (Same as 11, except that a key, piece of paper, and pencil are used, and that the objects are covered after being exposed to the subject for 30 seconds) Is the key to the left or to the right of the piece of paper? Same for all.

---

*Taken, with a few minor changes, from (3)

To answer all parts of the first six questions correctly, the subject must understand the inter-relationships of siblings both in the family of which he is a member, and in families of which he is not a member To answer the last six questions correctly, he must know that right and left for him do not necessarily coincide with right and left for somebody else, and that objects are not right and left in themselves, but in relation to other objects. Inability to abandon his own immediate point of view will prevent a subject from being able to answer these questions

In applying these tests to 240 school children ranging from four to twelve years in chronological age, and of unspecified mental age, Piaget found that there was a certain age level at which each test question could be answered by 75 per cent of the children, and, accordingly, a definite order of difficulty within the group of questions. The general problem of the present study has already been discussed; but the writers were also strongly interested in investigating to what extent these more specific findings of Piaget's would be duplicated in work with subnormal subjects.

The subjects were 50 subnormal children in Letchworth Village, 25 boys and 25 girls. Foundlings, only children, left-handed children, and those who had been in the institution longer than five years, were excluded The selection and grouping of the subjects were arranged to study the relative importance of mental maturity and chronological age This could have been done by holding the mental age of the subjects constant and varying chronological age, or by holding the chronological age constant and varying mental age, or by using a group selected for homogeneity as to one of the two factors and unselected as to the other. A complete study might make use of all three procedures; the present investigation used the second chiefly. In our groups mental age was varied, and chronological age held constant, for the most part, at a level where it is believed no longer to influence mental development markedly This was done by using four groups (Groups 1 to 4), of 10 children each, all having chronological ages of approximately sixteen, and mental ages on the Stanford-Binet Scale of six, eight, ten, and twelve years respectively. The average chronological ages and mental ages for these groups are shown in Table 1.

In 36 of the 40 cases in Groups 1 to 4, no chronological age deviated from sixteen years by more than four months. The remain-

TABLE 1

PERCENTAGE OF TOTAL NUMBER OF QUESTIONS ANSWERED CORRECTLY BY EACH GROUP*

Group	Average C.A.	Average M.A.	Test I (5 questions)	Test II (6 questions)	Both tests (11 questions)
1	16-6	5-11	10	25	18
2	16-3	8- 2	30	55	44
3	16-1	10- 1	78	80	79
4	16-2	12- 1	92	92	92
Y	12-1	8- 1	40	52	46

*N = 10 for each group.

ing deviations were only slightly greater. The *IQ*'s, of course, differed appreciably among Groups 1 to 4, Group 1 subjects having an average *IQ* of 38, Group 2 of 51, Group 3 of 64, and Group 4 of 77.

The fifth group, Group *Y*, was a contrasting group of 10 subjects who were chronologically younger, having chronological ages of approximately twelve years with no deviation greater than four months. Thus these subjects, as a group, differed equally in chronological age from each of the other Groups 1 to 4. But the subjects of Group *Y* had mental ages of approximately eight years (with no deviation greater than four months), which corresponded to the mental ages of the subjects of Group 2. Moreover, the chronological ages of Group *Y* corresponded to the mental ages of Group 4, while in *IQ* the subjects of Group Y corresponded to those of Group 3. The performance of these five groups of subjects, therefore, afforded opportunity for comparison independently on the basis of mental age, chronological age, and *IQ*.

C. RESULTS AND CONCLUSIONS

In studying the results we have examined both the total number of questions answered correctly and the difficulty of the individual questions. If chronological age determines the growth of relativism in these concepts, we should expect to find a considerable uniformity in the performance of the subjects in Groups 1 to 4, with an inferior performance on the part of the Group Y subjects. If, on the other hand, mental maturity is more important, we should expect to find that the performance would improve steadily from Groups 1 to 4, and that the performance of Group Y would more closely resemble that of Group 2 than that of any other group, since these contained subjects with the same mental ages

We may consider first the total number of questions answered correctly Here we wish to know how the series as a whole is handled by these children, and whether it discriminates between the groups

Table 1 shows the percentage of questions answered correctly by each group in Test I, comprising the questions on "brother and sister," in Test II, comprising the questions on "right and left," and in both series considered as a total Question 4, requiring the definition of a brother or sister, has been omitted here, for a reason which will be pointed out later.

It will be seen at once that there is a consistent and marked increase in the percentage of questions answered correctly in each test series, from Group 1 to Group 4 Thus, for our subjects, those with a mental age of six years answered 18 per cent of the questions correctly, those with a mental age of eight years answered 44 per cent, those with a mental age of ten, 79 per cent; and those with a mental age of twelve, 92 per cent Our data thus indicate clearly that in affecting the performance of these subjects mental maturity is a more important factor than chronological age

This conclusion is supported by an examination of the performance of Group *Y*. Here we find a percentage of correct responses which approximates the record of Group 2 far more closely than that of any other group, whether we consider each series of questions alone or both series together Group 2 subjects answered 48 questions correctly, or 44 per cent of the total; Group *Y* subjects answered 51 questions correctly, or 46 per cent of the total. The records of the two groups with the same mental age are thus found to be strikingly similar, whereas comparisons on the basis of chronological age or *IQ* show marked differences in performance.[4] These findings indicate, then, for these subjects, the existence of definite levels at which a certain number of correct responses can be made to the series of test questions considered as a battery, and they also bring out the importance of mental maturity in determining these levels.

There remains the problem of the comparative difficulty of the individual questions, and of the existence of levels at which specific questions can be answered. Table 2 shows the per cent of each group responding correctly to each question. Here we find that each question is answered, in general, by an increasing number of subjects as the mental level of the groups increases, and that questions answered at a certain level can always be answered at all higher levels. This means that each question in the series, like the series as a whole, discriminates to a certain extent between the groups.

---

[4] Comparison of the performance of Group *Y* subjects with that of Group 4, whose mental ages correspond to the Group *Y* chronological ages, reveals a particularly striking discrepancy

TABLE 2
PER CENT OF EACH GROUP RESPONDING CORRECTLY TO EACH QUESTION*

Question	Group 1	Group 2	Group 3	Group 4	Group Y
7	70%	100%	100%	100%	100%
9	40	90	100	90	80
2	20	60	90	90	80
8	20	70	90	100	60
10	20	60	90	100	60
3	0	40	100	100	40
5	10	20	80	100	30
6	10	10	70	100	30
11	0	0	50	80	10
1	0	10	40	70	10
12	0	0	50	80	0

*The questions are here arranged in order of decreasing number of correct responses from the whole group (50 subjects). Question 7 was answered by 94 per cent of all the subjects, questions 1 and 12 by 26 per cent. The one failure in Group 4 on question 9 was that of a boy who explained that he thought he should reverse his answer in this case to correspond to the examiner's point of view.

Our results also show, as did Piaget's, that the questions differ as to the levels at which they can be answered by a given percentage of the subjects. Table 3 indicates the levels at which the questions can first be answered, presenting these results in the form of a comparison between the Letchworth subjects and the Piaget subjects.

TABLE 3
LEVELS AT WHICH THE QUESTIONS CAN FIRST BE ANSWERED

	Questions answered by 70% of Letchworth group	Questions answered by 75% of Piaget group
6	7	2, 7
8	2, 8, 9	3, 8, 9, 10
10	3, 5, 6, 10	1, 5, 6
12	1, 11, 12	11, 12

In examining this table it must be borne in mind that the figures in the age column represent mental age levels for the Letchworth subjects and chronological age levels for the Piaget subjects Since the Letchworth results make use of only alternate years, these alone have been included here. One further point should be mentioned. whereas Piaget considered a test passed at a certain age if 75 per cent of the children of that age passed it, 70 per cent has been used as the criterion for the Letchworth groups, since each of these included 10 children

Looking first at the Letchworth results only, in Table 3, we may see that Question 7 is the only one which could be answered by a majority of the subjects in Group 1 This question, requiring the subject to name his own right and left hands and ears, involves no shift from the subject's own point of view in the response. Group 2 and Group *Y* subjects can answer, in addition to Question 7, Questions 2, 8, and 9 Question 2 involves simply knowing the total number of siblings in the subject's own family, and requires very little understanding of the reciprocal nature of the brother or sister relationship Because the total number of boys and girls in a family is frequently mentioned, it is doubtful that a subject must necessarily understand that he himself is a brother in order to pass this question To answer Question 9 the subject must be able to transfer the concepts of right and left from parts of his own body to objects outside of himself, right and left are thus no longer the names of simple qualities attaching to his own hands Question 8, the designating of the examiner's right and left hands, requires a still greater shift from the subject's own immediate point of view Question 10 necessitates a combination of these two processes, and is not answered until the next level; i.e., by Group 3 subjects The latter are also able to handle questions dealing with the relationships of brothers in families not their own, namely, Questions 3, 5, and 6 Not before the Group 4 level of twelve years can Question 1, regarding the relationships of the subject's own siblings be answered, and here it is answered by no more than 70 per cent of the group Question 1 provides considerable opportunity for confusion between the subject's own point of view and that of others The other questions answered for the first time at the Group 4 level, are 11 and 12, which deal with the positions of three objects, and requires a well-developed grasp of the relativism inherent in the concepts of right and left, for with these questions the subject must state that a key, for example, is to the left of a penny, but to the right of a pencil.

A brief comparison of the performance of the Letchworth group with that of the Piaget group shows that the range of levels, whether mental or chronological, is about the same for both groups. We may note especially that all of the questions can be answered in approximately the same order by both groups. They fall at approximately the same levels, with a few differences in favor of the Piaget subjects

Any work with Piaget's experimental techniques which deals solely with quantitative results neglects a part of the material which he himself emphasizes throughout. It is important not to consider the series of questions used here simply as a list of items which may be passed or failed. The content of the individual responses and the significance of the questions should be more thoroughly studied; and the reasons for failing certain questions analyzed In responding to Question 1, for example, it was common for a subject to fail for one of the following reasons. counting himself in the number of his own siblings, omitting himself in counting the brothers of his siblings, or saying that his siblings had no brothers or sisters These indicate a confusion on the part of the subject between the total number of children in his family and the number of his siblings, and between his point of view and that of the others Failures on Question 1 also included, particularly for the subjects of Groups 1 and 2, responses which were simply incoherent or self-contradictory. A few subjects who gave different numbers of siblings for different children in the family did so according to a plan, for example, attributing more brothers to the older children than to the younger Unfortunately, the limits of this paper do not permit more complete analysis of qualitative material. It should be pointed out, however, that Question 4 would be particularly useful in this connection This is the question which was not considered in the quantitative results because of the difficulty of scoring the responses as simply passed or failed Piaget stipulates that a successful definition of a brother must state that in order to be a brother one must have a brother or sister, but this criterion seemed to place too great a premium on a particular verbal form of describing the relationship.

The results of the study may now be summarized as follows: (*a*) The test questions used here, considered both separately and as a series, were handled progressively better by these subnormal subjects as the mental ages increased (*b*) Increase in chronological age alone, from twelve years to sixteen, did not improve performance (*c*) The test questions differ in difficulty, i e, in the mental level at which they can be correctly answered by a majority of these subjects (*d*) The levels found here correspond approximately to those found by Piaget for school children of different chronological ages Seventy per cent of our subnormal subjects with a mental age of six years answered only one of the test questions, whereas 70 per cent of those with a mental age of twelve years answered them all

The results indicate, therefore, that for mentally subnormal individuals, relativism in thinking, as measured by tests such as those used in this study, varies directly with mental level We may suggest more tentatively that, for any subjects, the ability to handle concepts involving such relativism depends more closely on mental level than on chronological age. If this is true it would mean that, when Piaget found the growth of relativism associated with increase in chronological age, the really significant relationship was rather between relativism and mental level.[5] It is clear that more work, using larger numbers of subjects, involving superior children as well as defectives, and varying both mental and chronological age at different levels, is needed to check any such general conclusion The rôle of social rather than intellectual maturity, in so far as this can be isolated, should also be examined.

The question of the extent to which it is legitimate to argue from results based on the particular test questions to conclusions regarding relativism and egocentricity in general, is one which cannot be ignored, although its treatment is outside the scope of this paper In this connection a study by T. M. Abel (1) and a recent article by L. Deshaies (2)[6] are of interest. M. Deshaies adduces experimental evidence to support the thesis that relativism is itself a "relative" concept, i e., that lack of relativism characterizes thinking in adults as well as in children, and that in its measurement the particular problems used to test it must be taken into account.

Whereas M Deshaies' study is directed toward a better understanding of the concept of relativism *per se,* the work described in the present paper seeks to test certain factors by which relativism may be influenced. Both approaches are important in clarifying Piaget's chief principles. A further important result of such work as this may be, as we have suggested earlier, a clearer understanding of other older and more familiar psychological concepts.

---

[5] Actually mental level and chronological age approximately coincide for the majority of school children

[6] This article did not come to the authors' attention until the present study was being prepared for publication

REFERENCES

1. ABEL, T. M. Unsynthetic modes of thinking among adults, a discussion of Piaget's concepts *Amer. J. Psychol.,* 1932, **44**, 123-132
2. DESHAIES, L. La notion de relation chez l'enfant. *J de Psychol ,* 1937, **33**, 112-133.
3. PIAGET, J. Judgment and Reasoning in the Child. New York: Harcourt, Brace, 1928.

Department of Philosophy
Yale University
New Haven, Connecticut

Research Department
Letchworth Village
Thiells, New York

The Journal of Genetic Psychology, 1939, 54, 119-149

THE EFFECTS OF COMPLETE AND PARTIAL OCCIPITAL LOBECTOMY UPON THRESHOLDS OF VISUAL REAL MOVEMENT DISCRIMINATION IN THE CAT*

Department of Psychology, Stanford University

JOHN L. KENNEDY[1]

A. INTRODUCTION

In the history of the growth of knowledge concerning mechanisms upon which behavior depends, the controlled preoperative and postoperative testing of animals after experimental lesions in the nervous system has been most important In applying this method to studies of receptor function, the point of view is now generally accepted that what is sometimes called the "perceptual" ability of different organisms may be studied and compared by establishing certain criteria of differential or selective behavior with regard to measureable differences in the physical environment The value of such an objective experimental method lies in the fact that it not only allows a rigorous control of motivation and suggestion, which is so often lacking in studies on humans after brain lesions, but it also permits the correlation of differences in functional capacity in so far as such differences may be found at various "levels" in the phyletic series

A large body of experimental data on typical laboratory mammals has been secured in the past 20 years bearing on the function of the cerebral cortex in the mediation of visually controlled behavior. These studies, although contributing many facts in regard to the rôle of the cortex in vision, have left unanswered some important

---

*Accepted for publication by Leonard Carmichael of the Editorial Board, and received in the Editorial Office on April 6, 1938.

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Brown University as a dissertation entitled "Visual movement discrimination in the cat. Effects of extirpation of the visual cortex upon movement discrimination"

[1] Fellow in Psychical Research, Psychology Department, Stanford University

The writer is indebted to Dr. K U Smith of the University of Rochester, who performed the operations on the animals used in the present experiment and aided in the histological work described

questions, among which is the problem of the function of the visual cortex in the perception of moving visual stimuli. The writer (15) has elsewhere discussed the available literature on this aspect of cortical function. The present report will be devoted to an analysis of residual capacity for movement discrimination in cats in which the visual cortex has been either partially or completely extirpated.

B THE ANATOMY OF THE VISUAL SYSTEM IN THE CAT

Before proceeding to the description of these experiments, it seems advantageous to present in detail the available anatomical data on the structure of the visual system in the cat. The importance of anatomical knowledge for psychological theory of the sensory mechanisms may be summarized in the words of Henschen (10) who stated,

> In order that modern psychology may possess a definite starting point and foundation, we must obtain an exact idea of the anatomical construction and physiological function of the paths and cortical areas of our senses, as well as their connections with other parts of the brain

The specific facts concerning the anatomy of the visual system of the cat can be assembled by combining experimental results from a number of sources. The following survey of the anatomical organization of vision in the cat is not complete, but it is presented in its present form to aid in evaluating the results herein discussed.

The exact anatomical aspects of the eye of the cat have been investigated by Ganser (7), reviewed by Vincent (44), and Smith (40) and completely presented, in comparison with the dog's eye, by Thieulin (42). According to the latter, the pupillary opening of the cat is approximately 16 mm at its maximum diameter. The curvature of the cornea is slightly less than in the dog and primates. The lens measures 12-13 mm in a transverse direction and is 8 mm thick. These factors and the degree of divergence of the optic axes determine the extent of the anatomical monocular visual field. In the cat the extent of the monocular field is 197-205 degrees, extending 105 degrees temporally and 100 degrees toward the nasal side. The binocular visual field extends 111 degrees. In comparison with the visual fields of other mammals, the binocular field of the cat is slightly larger than that of the dog but smaller than the binocular field of infra-human primates and man The monocular fields are approximately similar for the dog and cat. The total anatomical visual field is 297 degrees for the cat, slightly less for the dog.

The retina of the cat's eye has also been described by Thieulin (42) Vincent (44) quotes Krause as stating that the cat's retina contains a fovea but Thieulin presents definite evidence as to the absence of a true fovea An *Area centralis* with a maximum diameter of 1 mm which apparently contains cones only, is present. In a zone immediately surrounding the *Area centralis,* the cone-rod ratio is 1 20, more peripherally, 1 25 Thieulin was unable to detect visual purple in the cat's eye. Concerning the image-forming capacity of the lens and refracting media, it was found that the retina receives images of about three-fourths the size of the image on the human retina

The fibers forming the optic nerve of the cat originate in the outer ganglion cell layer of the retina At the chiasm, slightly more fibers cross over to terminate in the external geniculate body of the opposite side than continue on the same side The major portion of these crossed fibers carry impulses from the nasal halves of the two retinae, which receive light energy from the monocular portions of the visual field The temporal retinae, receiving light energy from the binocular portions of the visual field, send both crossed and uncrossed fibers to the external geniculate bodies [Minkowski (30)]

The lateral or external geniculate bodies are important nuclei of the diecephalon The optic tracts enter the brain on the ventral surface of the diecephalon and course upward to terminate in synaptic connection with cells in the dorsal nucleus of the external geniculate body. Some fibers pass through the geniculate bodies, running caudally, to terminate in and near the superior colliculi, phylogenetically old nuclei which are located on the dorsal surface of the mid-brain The connections of the retina with the superior colliculi and surrounding areas will be discussed later

According to the results of Minkowski (30), the whole visual field is represented in the lateral geniculate bodies, the portion to the right of the fixation point in the left external geniculate body and the left part of the field in the right external geniculate body Within each lateral geniculate body are found special projection areas for the monocular and binocular visual fields Minkowski's studies demonstrate that the retinae are projected on the external geniculate bodies in such a way that

> the lateral quarter of the retina on the same side is projected to the medial part (especially in the caudal part of the medial three-fourths) of the external geniculate body of the same side Concerning the remaining three-fourths of the retina, which is represented in the external geniculate body of the opposite side, the temporal quarter is projected in the medial portion of the caudal three-fourths, and the nasal half in the lateral part of the caudal three-fourths and the oral fourth of the crossed external geniculate body [2]

The more recent work of Overbosch (32) shows that in the external geniculate body the projection areas of the upper quadrants of the retinae lie frontally to the projection fields for the lower quadrants These results were obtained by producing small lesions in the retina itself and tracing the course of degenerated fibers to their termination in the external geniculate bodies.

The *Area centralis* of the cat which, as has been pointed out, seems to be the homologue of the *macula lutea* of primates, is the region of origin of both crossed and uncrossed fibers The exact localization of the fiber representation of this area in the external geniculate bodies is not known.

The schema of projection in the nuclei of the mid-brain, including the superior colliculi and the pre-tectal region, has never been fully worked out. Brouwer (2) presents evidence, based upon destruction of parts of the retina, for the existence of such projection in the rabbit and monkey. Lashley (21) has also found orderly projection of retinal fibers in the superior colliculi in the rat Historically, the superior colliculi have been considered to be centers mediating optic reflexes, such as pupillary constriction to light, through their connections with the motor nuclei of the 3rd cranial nerves Recent work on the cat by Ranson and Magoun (36), Magoun (24), and Keller and Stewart (12), with reference to the pathways mediating pupillary constriction, demonstrates that connections with the superior colliculi are not necessary for the maintenance of pupillary action to intense light stimulation. Fibers from the optic tracts terminate in the pretectal region as well as in the superior colliculi, and it is believed that these connections make possible pupillary action and eye-movements in the absence of the superior colliculus Ghiselli (8) has found that brightness and pattern vision in the rat is undisturbed after complete destruction of the superior colliculi Further studies on the function of the pretectal fibers of the visual system will be of great importance for theories of visual function. The mid-brain connections of the visual system of the cat are of special interest in the present paper since Marquis (26) has based a theory of encephalization of visual function in phylogeny upon the apparent shifting of function from the optic tectum (visual projection center of the mid-brain in birds and reptiles) to the geniculo-striate system of the cortex in mammals.

---

[2] Writer's free translation

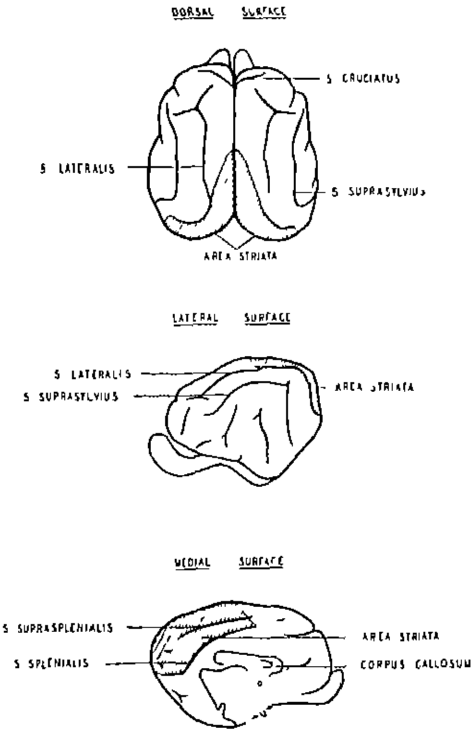


FIGURE 1
THE CEREBRAL CORTEX OF THE CAT

From the dorsal nuclei of the external geniculate bodies, second order optic fibers, forming the visual radiations, pass by way of the lateral saggital layer of the occipito-parietal lobe to terminate as stria Gennari-Vicq d'Azyr in the area striata of the cortex. Figure 1 shows the extent of the area striata on the dorsal and medial surfaces of the cat's brain This figure was prepared from similar diagrams presented by Campbell (4) In relation to the cell layers of the visual cortex, Poliak (35) concurs with the opinion of Brodmann that the majority of fibers of the visual radiation end in synaptic connection with cells in Layer IV, the *lamina granularis interna* of Brodmann.

The projection of the retinae on the visual cortex of the cat has not been completely determined, although Brouwer (2) reports that experiments are in progress to settle this question. Minkowski (29) concluded that the caudal portion of the area striata is connected with the temporal halves of the retinae, thus subserving binocular vision, while the oral part is similarly connected with the nasal retinae and the monocular visual fields

The relations of the afferent visual system, described above, are diagrammatically presented in Figure 2, which is based upon similar diagrams in the articles of Minkowski (29) and Barris (2). The connections to the pretectal region and the superior colliculus are shown only on the right side. The extent of the visual cortex on the medial surface of the left hemisphere and the projection of the external geniculate body upon the visual cortex is shown at the bottom of the figure

Barris (2) has recently summarized the available knowledge concerning the efferent fiber connections in the mid-brain of the cat. These fiber systems were traced by means of studying degenerations after enucleation of the eyes, by producing lesions in the dorsal nucleus of the external geniculate body and by destroying parts of the area striata. As mentioned previously, Minkowski (30) demonstrated that some afferent fibers from the retina terminate in the superior colliculus and the pretectal region The large efferent tracts arising in the superior colliculus and surrounding region, the tecto-bulbar and the tecto-spinal, are considered by Barris "to make possible the turning of the head and body in response to stimulation of retinal elements" He further states,

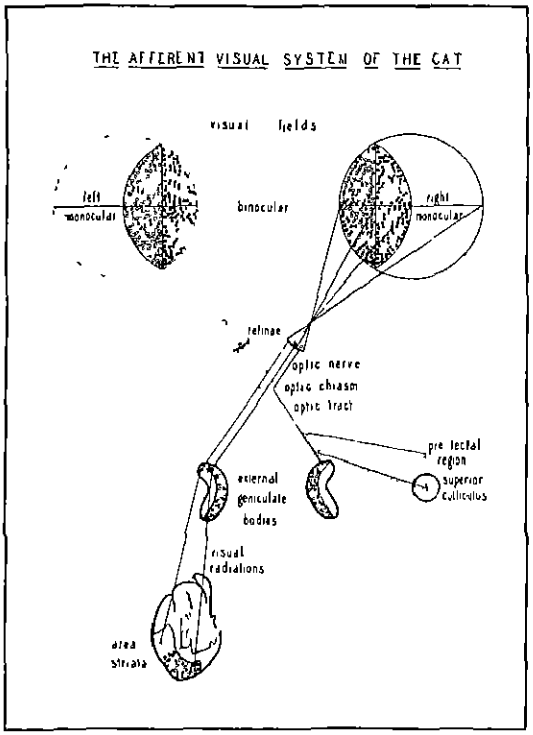


FIGURE 2
THE AFFERENT VISUAL SYSTEM OF THE CAT

> It is conceivable that the same impulse might reach the superior colliculus without first passing by way of the visual cortex Just what this means from a functional point of view is enigmatic, but that this possibility does exist seems sufficiently important to mention here

This statement will, perhaps, be clarified in the discussion of the results of the present experiments on extirpation of the visual cortex in cats.

Poliak's (34) (35) outstanding contributions on the neuroanatomy of the cortex of the cat furnish the basis for the following summary By tracing the course of degeneration after lesion in the visual area, he was able to ascertain that the striate and the parastriate areas send efferent fibers to the superior colliculus on the same side and to the nuclei of the pons by way of the internal capsule and the cerebral peduncle. Barris (2) has further shown that some of these fibers pass to the nucleus of the optic tract (*nucleus lentiformis mesencephali*), which lies just oral to the pretectal region between the diencephalon and the mesencephalon

As to the connections of the area striata within the cortex, Poliak (35) demonstrated that the cortex immediately surrounding it receives short fibers and in turn sends association fibers to it. He also states,

> The area striata is connected through callosal fibers with the opposite area striata and with the peri-parastriatal area or Field 18-19 there Homotopic region receives a greater number of callosal fibers, heterotopic, smaller

Thus it is evident that widespread changes due to fiber and cell degeneration take place in the central nervous system of the cat when the area striata is completely extirpated. In summary, these may be discussed as follows: (*a*) *Geniculo-striatal system*: Complete retrograde degeneration of the optic radiations and of the large cells in the dorsal nucleus of the external geniculate body; (*b*) *Tecto-collicular system*: Maintenance of functional mid-brain connections, concerning which little is known at present, (*c*) *Efferent system from the area striata*: Degeneration of fibers to the superior colliculus and perhaps to the pulvinar and other thalamic nuclei. Degeneration of fibers to the pons and the nucleus of the optic tract; (*d*) *Cortical system*: Degeneration of association fibers to Fields 18-19 of the same side and callosal fibers (passing through the splenium) to Fields 17, 18 and 19 of the other side.

C. PRESENT INVESTIGATIONS ON THE EFFECTS OF REMOVAL OF THE VISUAL CORTEX UPON MOVEMENT DISCRIMINATION

1. *Problems*

The problems on which the results of the experiments reported here have a bearing may be stated as follows:

(1) What is the threshold capacity in normal cats for the discrimination of: (*a*) moving from non-moving identical patterns, (*b*) the faster of two moving identical patterns?

(2) How much do the lower absolute movement thresholds change after partial bilateral extirpation of the area striata?

(3) What effect does complete bilateral extirpation of the visual area have upon absolute movement discrimination thresholds?

(4) What conditions will produce significant discriminatory responses to simple pattern differences after complete loss of the visual cortex?

(5) What effect does extirpation of the visual area have upon simple localization of brightness contours?

2. *Apparatus*

The apparatus used in these experiments has been described previously by Smith and the present writer (13) Figure 3 presents a side view of the apparatus as a whole. Two lever discrimination boxes (*B*), of the sort described by Smith (37), are so arranged on a long table that various illuminated patterns can be projected on diffusion screens in the front doors of the boxes. The box is constructed in such a way that the depressing of a lever (*L*), which projects from the front face of the box, releases a spring catch attached to the lever. The releasing of this catch allows the door to open and permits the animal to gain access to the interior of the box where food (small pieces of salmon) is placed before each trial.

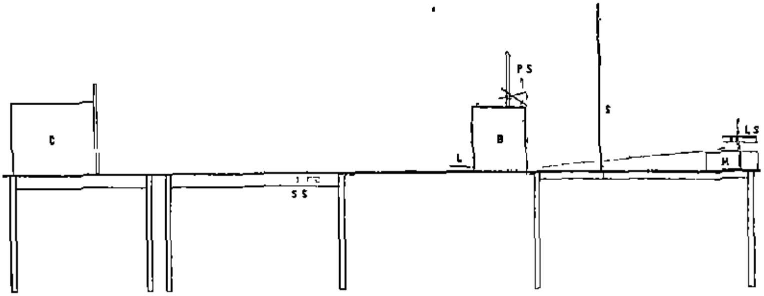


FIGURE 3
SIDE VIEW OF THE APPARATUS
C—Restraining Cage, *SS*—Shocking Switches, *L*—Lever, *B*—Discrimination Boxes, *PS*—Pulley System, *S*—Screen, *M*—Phonograph Motor, *LS*—Light Source

The door of the box in which the negative stimulus is presented is locked with a thumb latch so that the depressing of the lever does not open the door.

Small, specially devised projection lanterns (*LS*) are used as light sources. These are constructed from a length of copper tubing, a 32 candle power headlight bulb and an eye-piece of a common compound microscope. They are wired in series and receive current from a 12 volt transformer. With these light sources, an equal illumination of approximately 2 apparent foot candles was maintained on both screens throughout the experiments. A large cardboard screen (*S*) effectively eliminates all the light except that which illuminates the patterns on the doors of the boxes.

The response required from the animal in this situation is as follows. Upon release from the restraining cage (*C*), it is trained to run toward the stimuli and to choose between an arbitrarily selected positive and negative visual pattern by depressing the lever just below the positive stimulus. Successful completion of this spatial discrimination is rewarded by food If the animal chooses the negative stimulus it receives mild electrical punishment of known amperage through electrodes consisting of the plate of the lever itself and a copper screen on the table The control of the position of the shock was effected through two switches (*SS*). Several types of shock indicator were used during the experiments The most satisfactory one consisted of a neon bulb, wired in series with the shocking circuit, which flashed when the animal completed the circuit. The use of these bulbs was suggested by the similar system described by Henry (9)

Figure 4 shows the backs of the discrimination boxes (*B*) and the method of producing two different velocities of revolution of the patterns Light from two sources is projected through the glass plates (*GP*) which form the backs of the boxes and through the cardboard stimulus cards (*S*) to the screens on the front doors of the boxes

In order to make the stimulus cards movable, a small hole is drilled in the centers of the glass plates. The glass serves as a support for the brass bearings of a small steel shaft. Outside the box a two-stage pulley is attached to the shaft. Inside, the cardboard stimulus cards are attached as shown in Figure 4. Thread connections to a reduction pulley system, *P*1, *P*2, *P*3, and *P*4, and thence

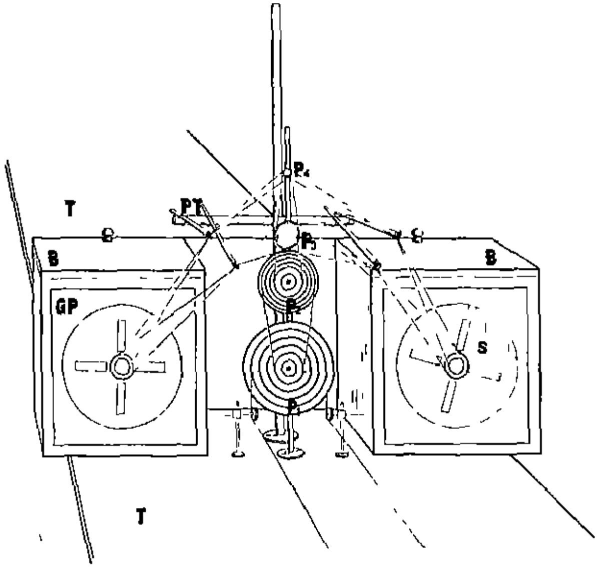


FIGURE 4

Backs of the Discrimination Boxes

T—Table; *B*—Discrimination Boxes, *P*—Glass Plate, *PT*—Pulley Tighteners, P_1, P_2, P_3, and P_4—Reduction Pulley System, *S*—Stimulus Card

to an electric phonograph motor (*M* in Figure 3) make possible the rotation of these patterns at any desired speed throughout a wide range. Rotation may be started or stopped by silently setting small pulley-tighteners (*PT*).

In finding absolute movement thresholds, as described previously (13), only the lower set of pulley tighteners was used since only one of the pair of stimuli revolved In order to rotate the figures at two different speeds when determining relative movement thresholds, a standard speed is obtained in both boxes by means of two identical stages on Pulley 3 and the lower two pulley-tighteners Either a faster or a slower speed, depending upon the pulley ratios utilized, may be obtained with two identical stages of Pulley 4 and the upper pulley-tighteners. To operate, the lower pulley-tighteners are both set to give the standard speed when the upper ones are loosened, thus enabling the experimenter to change the speed of either stimulus by tightening or loosening the upper pulley-tighteners only. The thread connections from the pulleys slip in the pulley stages without affecting the rotation of the stimuli when the tension on them is reduced to the correct degree Repeated tests show that no differential sound cues are available to the animal in this situation since the apparatus is almost noiseless in operation and since constant procedure is adhered to in preparing the animal and the apparatus for each trial.

Figure 5 presents a diagram of the basic patterns as they were

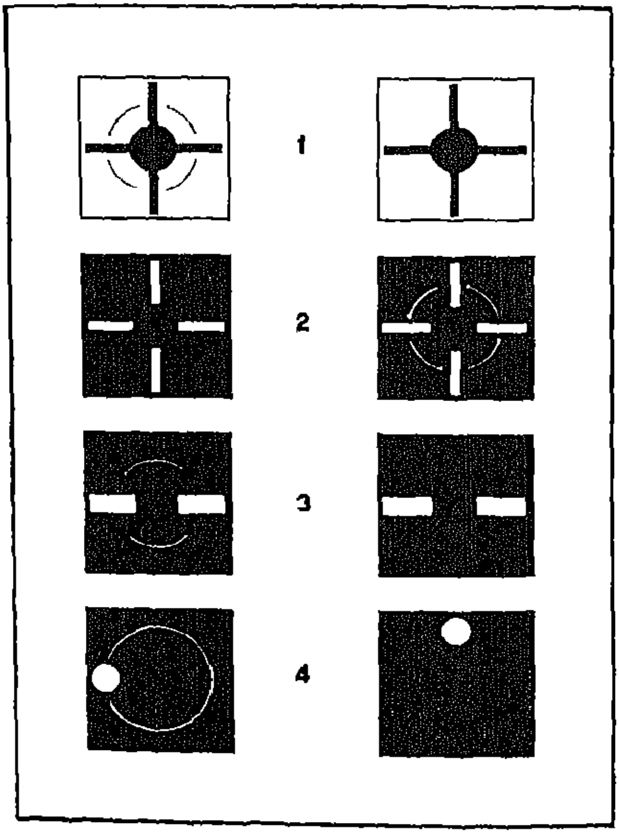


FIGURE 5
STIMULUS COMBINATIONS

shown to the animal Stimulus Combination 1 consisted of shadows of crosses, 9.5 cm in radius, with arms 1.5 cm thick, situated on illuminated backgrounds which were maintained at two apparent foot-candles brightness Stimulus Combination 2 involved merely a reversal of the brightness values of figure and background Stimulus Combinations 3 and 4 maintained the brightness relations of Combination 2, but the brightness gradients were fewer and less complex The manner in which these stimuli were used will be described in conjunction with the report of the experimental procedure and results. They will be designated by number hereafter

3 *Subjects and Procedure*

Visual movement thresholds were found with 10 normal cats These animals were selected from a larger group for general health and amenability to the training procedure They were maintained in excellent health and were hungry enough at the start of each experimental session to perform consistently in the trial series. In training the animals to perform the necessary discriminations, varying periods of pre-training, in which the subject learned to depress the lever and obtain food in either box, were given When this lever habit had been thoroughly mastered, all the cats were trained to a criterion of not more than four choices of the negative stimulus in 40 trials. Twenty trials per night were required of each animal. The position of the positive and negative stimuli was changed in accordance with a modified chance order, consisting of both double and single alternations, so that the animal was forced to depend upon visual cues only. Threshold determination proceeded from this point in the training.

4 *Operations and Anatomical Controls*

At the conclusion of pre-operative threshold determination the cats of the present study were subjected to an aseptic operation in which both visual areas of the cerebral hemispheres were removed in one stage. The animals were anesthetized with Nembutal, Veterinary, 75 cc per kilogram body weight. Entrance through the skull under the temporal muscle was effected by means of a small electrical burr. The bone over the occipital lobes was removed with bone forceps and the dura cut and reflected With Subjects 32, 33 and 34, the visual cortex was injured with a thermocautery, producing partial lesions in the monocular and binocular areas. In the remaining seven subjects the occipital lobes were dissected out in their entirety by means of the cautery, spatula, and bent-nosed forceps. Bleeding from the bone was checked with aseptic wax. After hemorrhage in the cortex was stopped with small cotton pads soaked in salt solution and the clotted blood fully removed, the wound was closed by suturing the temporal muscles, the superficial muscles and the skin The animal was allowed two weeks for recovery before it was tested and retrained under the same conditions which were utilized before the operation. The health of the animals was uniformly excellent throughout post-operative tests. Only one, Subject 34, developed a serious infection.

When post-operative testing was terminated, the brains of the animals were perfused with formalin through the carotid arteries They were then carefully removed from the skull and hardened in a series of alcohols Camera lucida drawings were made of the extent of destruction in the visual area. Photographs of the brains of Subjects 36, 37 and 38 are shown in Figure 6 as examples of complete removal of the area striata. The brains of Subjects 17, 32, 33, 34, 39, 40, and 41, were embedded in celloidin and serial frontal sections, 60 micra in thickness, were cut through the entire cortex. Every second section at the level of the external geniculate bodies and every tenth throughout the rest of the cortex was stained with thionin. The brains of Subjects 36, 37, and 38, were cut in blocs including the lesion, embedded in parafine and serial sections, 20 micra thick, were cut on a plane perpendicular to the brain stem to facilitate comparison with the Winkler and Potter series of the normal cat's brain (45). Every tenth section was stained with thionin throughout the bloc.[8] Figure 7, showing graphical reconstructions of the extent of destruction in the brains of these animals, was prepared from examination of the excised blocs of tissue which were fixed at operation. These reconstructions were later corrected on the basis of drawings made of typical sections through the lesion in each brain. The brains of animals with complete bilateral operations, Subjects 17, 36, 37, 38, 39, 40, and 41, were examined for

---

[8] This series was prepared through the courtesy of Prof H S Burr at the Section of Neuroanatomy, Yale School of Medicine, and was financed by a grant from the Psychology Laboratory, University of Rochester

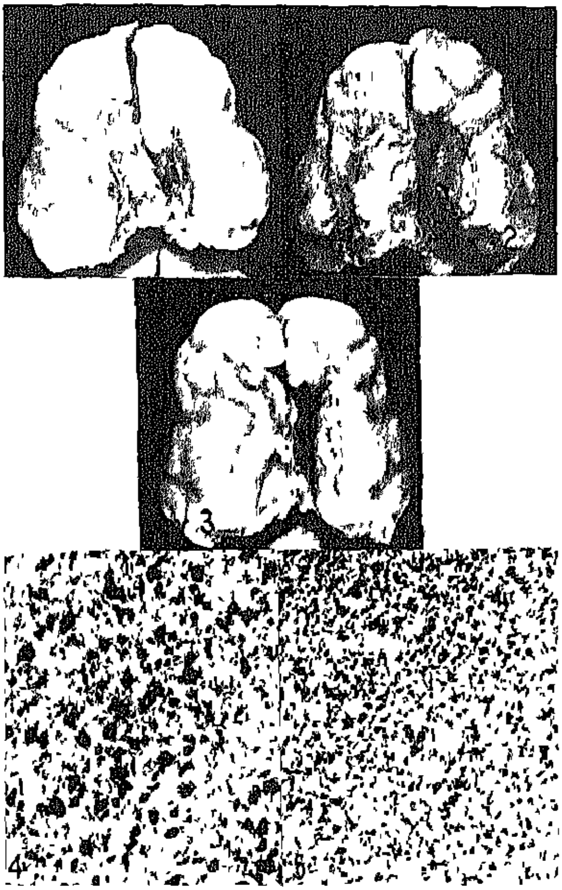


FIGURE 6

1. Cortex of Subject 36
2. Cortex of Subject 37
3. Cortex of Subject 38.
4. Photomicrograph of a portion of a normal dorsal nucleus of the external geniculate body of the cat (180x)
5. Neighboring portion of degenerated dorsal nucleus. These two photographs were taken from a section, cut at 60 micra and stained with thionin, of the brain of a cat with a partial lesion in the left area striata. The dorsal nuclei of Subjects 17, 36, 37, 38, 39, 40, and 41, were completely degenerated, presenting the appearance of 5

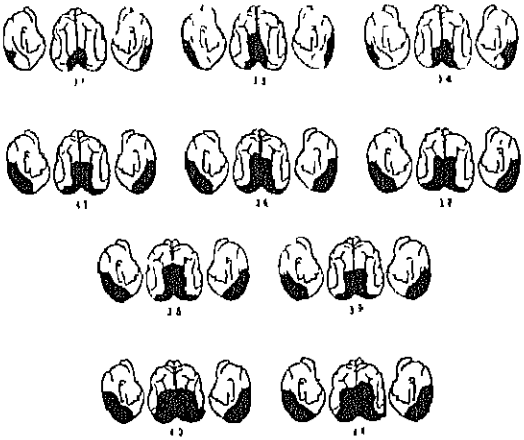


FIGURE 7

RECONSTRUCTIONS OF LESIONS IN THE BRAINS OF THE EXPERIMENTAL ANIMALS
(Black areas denote extirpation of tissue at operation)

the degeneration in the dorsal nucleus of the external geniculate body which has been described by Minkowski (29), (30), Brouwer (2), Poliak (35), and others. Figure 6 also shows photomicrographs, taken at 180 diameters, of typical normal and degenerated areas in the dorsal nucleus. The presence of complete degeneration in the dorsal nucleus of the external geniculate body in the animals with complete bilateral lesions lends additional proof that all cells in the area striata were either removed or that the visual radiations were undercut by the operation. Further examination of the mid-brain in the slides of Subject 41's brain showed that a small lesion had been made in the region of the superior colliculi at operation.

D. RESULTS

1. *Threshold Capacity in Normal Cats for the Discrimination of Movement*

After Subjects 17, 32, 33, 36, and 37, had fulfilled the criterion of learning in the initial training to discriminate between a cross moving at 80 degrees of central angle per second and a stationary cross, the speed of the revolving cross was reduced in small gradations. The subjects were tested at each speed until a significant level of discrimination was reached. Lower absolute thresholds, obtained by selecting the slowest speed of the moving cross which the animals could differentiate in comparison with the stationary cross, were respectively, 2.6, 5.5, 14.6, 3.0, 14.6, and 5.5 degrees of central angle per second. Two thresholds were determined with Subject 17. Stimulus Combination 1 was utilized with Subjects 32, 33, and 34, Combinations 1 and 2 with Subject 17 and Combination 2 with Subjects 36 and 37. A detailed account of the procedure is contained in reference (13).

Similarly, immediately after Subjects 38, 39, 40, and 41, had reached a significant level of response in the initial movement-no-movement situation, the stationary cross was set moving at a velocity of 11.25 degrees per second (ratio of 8 with the standard) and subsequently raised in speed in successive trial periods until the velocity of the standard was reached. The animal was required to react positively to the standard speed of 80 degrees per second. In Table 1 are shown the speed ratio series, the best 20 trial period at each ratio for each animal and the magnitude of the threshold ratios obtained with the four cats. The threshold of relative movement discrimination is here defined as that speed ratio at which the animal makes 75 per cent correct responses or better in a 20-trial period and below which its performance drops to a chance level consistently. Two thresholds were determined with Subject 38 in order to check on the

TABLE 1

RELATIVE MOVEMENT THRESHOLDS

Standard Speed: One Revolution in Four Seconds

Speed Ratio	% Correct choice in best 20-trial period				
	Subject 38		Subject 39	Subject 40	Subject 41
8.00	90	90	95	95	95
6.25	100	100	90	100	100
5.00	100	100	100	95	100
3.75	90	95	95	100	100
3.00	90	95	95	90	95
2.50	85	95	100	85	90
2.00	85	95	90		90
1.50	75*	90	85	90*	90*
1.25		95*	75*	70	70
1.00		50	45	65	50

* Indicates threshold ratio.

reliability of the training method Speed ratios of 1.20 and 1.15 were also tried with the same subject on the second determination, but repeated tests on the speed constancy of the apparatus showed that such small differences could not be reproduced with sufficient accuracy to warrant further work. As presented in Table 1, the limits of discriminatory capacity of the different animals with respect to differences in velocity between two moving visual patterns are, respectively, speed ratios of 1.25, 1.25, 1.5 and 1.5 with the standard Stimulus Combination 2 was used throughout these tests.

2 *Modifications of Lower Absolute Movement Thresholds after Partial Bilateral Lesions in the Visual Area*

Table 2 shows the course of post-operative threshold determination as contrasted with pre-operative thresholds in the three animals, Subjects 32, 33, and 34, in which the destruction of the area striata was definitely sub-total. The extent of change in visual capacity for

TABLE 2

The Effect of Partial Bilateral Extirpation of the Visual Cortex on Lower Absolute Thresholds of Real Movement Discrimination

Speed of cross (degrees/sec)	Per cent correct choice of moving cross in best 20-trial period					
	Animal 32		Animal 33		Animal 34	
	Pre	Post	Pre	Post	Pre	Post
80.0	100	95	90	90	90	100
33.6	95	85	100	80*	80	100
30.0	90	90	85	70	85	95*
24.8	85	75*	75	60	60	
21.1	90	70			50	
14.6	85	55	90		75*	
11.2	75				55	
6.5	90		80		40	
5.5	75		75*			
4.3	75		55			
3.7	75		55			
3.4	65					
3.1	80					
3.0	70					
2.9	85					
2.6	85*					
2.4	60					
2.2	60					
Total trials	1220	620	900	560	920	800

* Indicates threshold speed

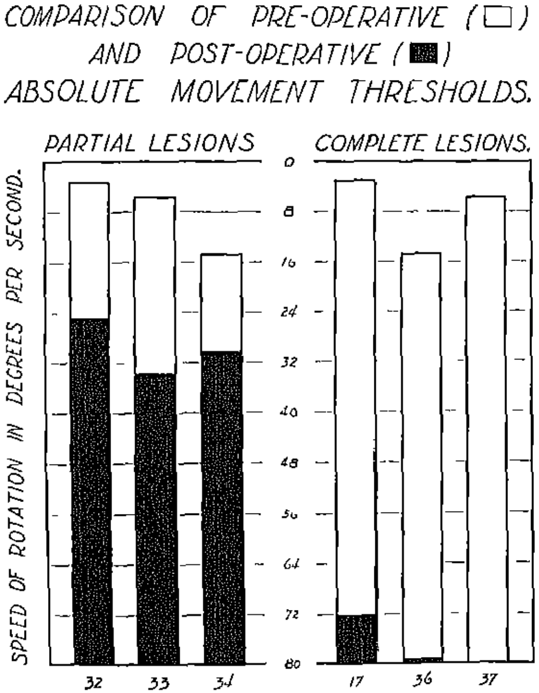


FIGURE 8

movement discrimination is also presented graphically in Figure 8. These data demonstrate that partial bilateral destruction of the striate areas produces a measurable decrement in the ability to differentiate between moving and stationary visual patterns The three animals in this series began to make discriminatory responses immediately after they were returned to the apparatus following the operation. Stimulus Combination 1 was used both pre- and post-operatively in finding these changes in threshold

3. *The Effect of Complete Lesions on Lower Absolute Movement Thresholds*

The animals with total removal of the area striata were given 1000 to 2000 post-operative trials and were observed in some cases for over a year. This period of testing afforded opportunity for making observations on residual vision using a number of different arrangements of basic patterns and movement speeds. After working with the animals for 500 trials, it became evident that post-operative threshold determination with stimulus Combinations 1 and 2 would be very difficult due to the extreme variability of performance from trial period to trial period and even within a single experimental session The patterns of Combinations 3 and 4 were then added to the training series in an attempt to determine if less complex brightness relations or patterns might not be more efficacious in producing discriminatory responses. In all cases when movement vision was tested, the animal was required to choose the revolving stimulus from a choice between a moving and a stationary identical pattern Relative movement discrimination was not attempted because the animals were unable to maintain a significant level of response with stimulus Combination 2. A great deal of difficulty was encountered in forcing the animals to use visual instead of kinesthetic and tactual cues. Consequently, position habits occurred with great regularity and in an unpredictable way. These methods of response were effectively discouraged by increasing the hunger drive, by severe electrical punishment, and by aid from the experimenter. In some cases, shocking the animal produced an "experimental neurosis," during which it refused to approach the discrimination boxes Reduction of the shock and increasing the drive invariably "extinguished" this conditioned reaction. Thus, throughout the post-operative training period, the experimenter worked in that difficult situation of position habits on the one hand and experimental neuroses on the other.

As indicated above, the determination of post-operative absolute thresholds proved to be a difficult task for the animals. The variability of performance of the animals after total extirpation of the visual area precluded any detailed and valid threshold measures. Figure 8 graphically presents a comparison of the pre- and post-operative absolute movement thresholds for three of the animals. Subject 17 showed definite evidence of discrimination at the original speed of the moving cross of 80 degrees of central angle per second and made a score of 75 per cent in one trial period at 72 degrees per second Hence, the latter speed was chosen as expressing roughly the limits of the animal's visual capacity for these stimuli. Subject 36 also gave evidence of discrimination at 80 degrees per second, but its performance did not warrant any further speed reduction Both animals discriminated well at a speed of the moving cross of 200 degrees per second, a speed far out of the range tested pre-operatively. As Figure 8 shows, Subject 37 failed to discriminate with stimulus Combination 2 at any speed Visual function for other patterns was not completely lost, however.

4. *The Performance of Cats with Complete Occipital Lobectomies in Movement Discrimination with Simple Figures*

The rather equivocal level of response obtained with stimulus Combinations 1 and 2 suggested that a more significant series of discriminations might be obtained with patterns in which the intensity relations are less complex Combinations 3 and 4 were therefore selected for further testing. This hypothesis proved to be true, as reference to Table 3 will show. Table 3 presents the significant per cent correct choice of the moving stimulus in the best 20-trial period with Combinations 2, 3 and 4, as well as the number of post-operative trials preceding this trial period The speed of the revolving stimulus in degrees of central angle per second and the position of the stationary stimulus is also given. As shown in this table, Subject 39 also reached a significant level of response with Combination 2 but the animal died before its sensitivity to differences in velocity could be tested. The performance of Subjects 40 and 41 was uniformly poor, but they succeeded in discriminating both Combinations 3 and 4 at relatively high speeds of the moving stimulus.

During the course of experimentation with Combinations 3 and 4, some results were obtained which appear to bear upon the problem of stationary pattern discrimination in the cat after extirpation of the visual cortex Observations were made on the ability of these animals to respond discriminatively to a moving stimulus when the position of the stationary or negative stimulus was changed In the case of Subject 17, change in the position of the stationary spot (Combination 4) significantly modified the immediately preceding level of response with the spot at the top of the stimulus panel Similar results were obtained with the other animals Although such evidence does not prove the existence of stationary pattern vision of a simple order in cats without visual cortex, it does suggest the possibility that the stimulus conditions of the present experiments might produce differential responses to differences in stationary patterns This suggestion seems all the more likely since Koudrin (18) has reported the elicitation of conditioned salivary responses to simple pattern stimuli under essentially the same conditions in a dog with complete occipital lobectomy

TABLE 3

THE PERFORMANCE OF CATS WITH COMPLETE BILATERAL EXTIRPATION OF THE VISUAL CORTEX IN DISCRIMINATING MOVING VISUAL PATTERNS

Stimuli	Speed of moving stimulus	Position of stationary stimulus	Subject 17		Subject 36		Subject 37		Subject 38		Subject 39		Subject 40		Subject 41	
			%	Tr	%	Tr.	%	Tr	%	Tr	%	Tr.	%	Tr.	%	Tr.
2	72 °/sec	Chance	75	1080												
	80 °/sec	Chance	85	1060	80	340					80	680				
	200 °/sec.	Chance	85	980	85	1040					85	940				
3	80 °/sec	Horiz					80	360								
	200 °/sec.	Horiz	95	920	90	1020			60	380	90	900	80	1280	75	1080
	200 °/sec	Vert	90	940												
4	200 °/sec.	Top	90	480	90	920	75	1040	90	820	90	980	85	140	75	460
	200 °/sec.	Bottom	90	880	75	1420			80	1320	95	840				

5 *The Effect of Complete Lesions upon Responses to a Single Moving Stimulus*

Subjects 17, 36, 37, 38, and 41, were tested for retained capacity for differentiating between a moving pattern on the screen of one box and the uniform black screen of the other. In the case of Subject 36, the first presentation of these stimuli brought about perfect discrimination, whereas in 480 previous post-operative trials in a two-stimulus situation, where the gross brightnesses of the two panels were equated, its performance was extremely variable and generally poor Essentially the same results were obtained with Subjects 17, 37, and 41. Subject 38 was started post-operatively on this brightness discrimination problem to determine whether or not the animal could perform consistently immediately after the operation. Significant discrimination was not produced until the 13th trial period or after 240 trials

Detailed protocol records of the performance of the animals were kept for each animal in each trial period [4] In comparison with the performance of the normal animal, the variability exhibited by the operated cat in the same situation is large. Apparently the general health of the animal is one factor with which such variability may be correlated It may even be that trophic variation in the condition of the nervous system should be looked upon as a possible explanation. On the other hand, the variable behavior found in these animals is characteristic of the normal cat in a situation in which the cues for discrimination are greatly reduced The problems in movement discrimination set for the operated animals seemed to be definitely of this sort

E. Summary and Discussion

The results of the present experiments may be discussed and summarized under the following topics: (1) the results of experimentation with normal animals; (2) the effects of partial bilateral destruction of the area striata on the visual capacity of the cat, (3) the nature of losses in visual capacity after complete bilateral extirpation of the visual area of the cortex in the cat, the relation of these results to theories of cortical visual function and the role of the cortex in movement perception

---

[4] These summaries of daily records are obtainable in the dissertation on file in the Brown University Library

1. *The Results of Experimentation with Normal Animals*

The lower absolute thresholds of visual movement discrimination obtained with some of these animals have been discussed at length elsewhere (14), (15), in relation to studies on other mammals and man in which similar thresholds were found A summary of literature on relative movement discrimination is also presented in reference (15). The results obtained here on relative movement discrimination are consistent with the general high level of visual sensitivity of the cat as shown in the previous experiment on absolute movement thresholds and by the visual acuity results reported by Smith (39). An accurate comparison of the data presented here and those of other experiments on animals in which movement thresholds were found is difficult to make since the stimulus conditions of all these experiments differed markedly. In general, it may be said that, with respect to the discrimination of moving visual stimuli, vision in the cat is organized on a level but slightly lower than that of primates.

2 *The Effects of Partial Bilateral Destruction of the Area Striata on the Visual Capacity of the Cat*

The animals with partial lesions gave the impression of having some residual visual capacity under the illumination conditions of the living quarters. Bumping into the wire of the cage and into other objects such as food pans was not noticed as frequently as with animals having complete lesions. Pupils were normal after the recovery period and reaction to change in illumination seemed to be at normal speed. The general activity level was lowered in Subjects 32 and 33, while Subject 34 was slightly more active than the normal cat, due probably to the development of infection. Numerous observations at feeding time show that the ability to find food when it was placed in the cage was only slightly impaired in these animals The results on cats with partial lesions in the area striata fit into the picture of other mammals with partial unilateral and bilateral lesions which has been developed through the work of Munk (31), Loeb (23), Minkowski (29), Lashley (19), (20), (21), (22), Smith (38), Penard (33), Kluver (16), and Spence and Fulton (41), with the dog, cat, rat, monkey, and chimpanzee. A small decrement in all visual functions appears to result from either partial unilateral or partial bilateral extirpation in the area striata. According to anatomical studies on the visual system of the cat, the lesions in the present series destroyed parts of both the binocular and monocular cortical areas The perimetric experiments of Minkowski on kittens suggest that partial lesions of the area striata produce limited "scotomata" or regions of blindness A moving visual stimulus is probably more efficacious than a stationary one in stimulating the normal parts of the retina in such a preparation.

In general, a moving visual pattern appears to be more effective in the determination of residual visual function after cerebral injury in animals and man than any other type of "patterned" visual stimulus. In this connection, it is interesting to note that Holmes (11), in discussing some human cases of partial lesion in the occipital lobe says, "only large white objects when moving may be recognized . . when they cease to move, he sees them no longer, they disappear "

3. *The Nature of Losses in Visual Capacity in the Cat after Complete Bilateral Extirpation of the Area Striata*

The seven animals with total bilateral lesions appeared to be completely "blind" when observed in the living quarters under conditions of strong daylight illumination Their general activity was at a very low level in comparison with the normal state These animals were able to find food placed in their cages only if adequate auditory and tactual cues were given They learned to find food if it was always placed in the same position They would not descend from platforms, visual placing was absent, and falling from the table of the apparatus and from the living cages was a frequent occurrence

However, it does seem possible to conclude that the animals with complete lesions were capable of making certain responses which can be defined objectively as "movement" or "pattern" vision The responses of Subject 17, for example, in the post-operative attempts to determine a lower absolute movement threshold with stimulus Combination 2 gave some significant evidence of discrimination based upon the use of visual cues. Figure 8 shows that although practically all "useful vision" for these stimuli had been abolished by the operation, the animal was still able to perform consistently at high speeds of movement of the positive stimulus Subjects 36 and 39 gave similar results. All subjects, with the exception of 40 and 41, made some significant responses with the less complex stimulus patterns and higher speeds. Subject 41's performance may be explained by damage to the mesencephalic connections of the visual system as well as complete extirpation of the visual cortex

Some of the animals of the present study were used by Smith (40) in studies of visual placing reactions of the forelimbs, pupillary and palpebral responses to sudden changes in illumination, and eye nystagmus in a rotating visual field. No visual placing reactions could be obtained with Subjects 36, 37, and 38. Pupillary constriction and palpebral closure, however, were elicited under several conditions of change in illumination The nystagmus responses of these subjects in a rotating visual field, consisting of alternate black and white stripes 1 in wide, could not be differentiated from those of normal cats. Smith concludes, and the results of the present study also definitely indicate, that these animals cannot properly be called blind to patterned visual stimuli, although it is evident that their localizing capacities are greatly reduced

F. THEORETICAL

The effects of total removal of the area striata in representative laboratory mammals on visually-controlled responses has been described previously by Lashley (19), (20), (21), (22) for the rat, Van Herk and Ten Cate (43) for the rabbit, Marquis (25), Marquis and Hilgard (27), and Koudrin (18) for the dog, Smith (38), (40), and Ten Cate (45) for the cat, and Kluver (16), (17), and Marquis and Hilgard (28), for the monkey Marquis (25) has also reviewed the findings on nine human patients with complete bilateral lesion of the area striata. The general rule has been accepted that the debilitating results of this serious loss of cortical tissue increase markedly as the phylogenetic scale is ascended According to the pre- and post-operative comparisons of Lashley, Marquis, and Smith, the ability to discriminate brightness differences, using transmitted light under conditions of general low illumination, is not significantly changed Lashley, however, was unable to obtain discriminatory responses based upon differences in simple stationary or moving pattern stimuli under conditions of high or daylight general illumination with the rat Observations of other animals with the visual cortex removed, under the latter conditions of illumination, have all been consistent in demonstrating that no differential responses to patterned visual stimuli can be elicited It has been generally concluded from these studies that, after complete extirpation of the visual cortex in mammals (with the exception of man and perhaps the anthropoid apes), elementary brightness discrimination is retained while "pattern vision" or "object vision" of any sort is lost

The experimental facts derived from previous studies on the extirpation of the visual cortex in mammals, from clinical observations on man and from the comparative neuroanatomy of the visual system have been summarized by Marquis (25), (26), and Dusser de Barenne (6) These authors have presented a theory of progressive "encephalization" or "corticalization" of visual function in phylogeny Marquis states,

> It is thus apparent that in the evolutionary development from fish to man a complete reorganization of the central visual system has taken place In the lower vertebrates, the visual functions are carried out entirely by the mesencephalic centers, while in man the cortex has taken over all optic functions with the exception of the pupillary reflex to light

As to the function of the visual cortex in mammals, Marquis summarizes experiments on monkeys and dogs in the following words

> Its functions are those of fine discrimination and of spatial perception, as shown by the impairment of the threshold of sensitivity to light and the complete loss of vision for objects

Certainly, the *qualitative* observations presented here on cats without visual cortex are, in general, in agreement with this statement The choice of the terms "brightness vision" and "pattern" or "object" vision in describing the *quantitative* losses occasioned by total extirpation of the area striata, however, is perhaps unfortunate in that they apparently refer, in the case of brightness vision, to the results of experiments conducted under conditions of low general illumination of the retina and, in the case of pattern or object vision, to experiments under conditions of high retinal illumination The results of the present experiments indicate that for the cat, at least, *complete* loss of movement or pattern vision, defining a visual pattern as a more or less complex concatenation of brightness ratios or gradients, does not occur. It appears that the loss of the visual area in the cat causes a huge decrement, but not a total loss in ability to discriminate between stationary and moving patterns having relatively small contrast differences. It is the present writer's belief that conditions of high contrast between figure and background, sharp contour and more complete dark-adaptation could be found under which the cat lacking the visual cortex would show little or no decrement when compared with its normal performance.

G CONCLUSIONS

1. These experiments were designed to attempt to quantify the effects of partial and complete extirpation of the area striata of the cat upon its known capacity for discriminating, by means of differential turning responses of the whole body and visual placing of the fore limbs, between revolving and stationary visual patterns.

2. Two thresholds of visual movement discrimination were determined prior to the operation · (*a*) lower absolute thresholds, found by reducing the speed of a moving cross until the animal was unable to maintain a significant discrimination between the moving and an identical stationary cross, (*b*) relative thresholds, determined by increasing the speed of an originally stationary cross to approach that of a standard revolving cross.

3. Three animals of the absolute threshold group were subjected to partial bilateral extirpation of the visual cortex. Their postoperative capacity was significantly but not greatly reduced.

4. Complete bilateral extirpations of the visual area were performed on seven cats, one of which had additional lesions in the superior colliculi. Post-operative lower absolute movement thresholds were found with two animals, but variability of behavior makes these results difficult to interpret. These thresholds were greatly changed, corresponding closely to a total loss of visual capacity under the conditions of the experiment. Relative thresholds were not found postoperatively since the subjects were unable to maintain discrimination at a significant level with the pre-operative patterns When simple patterns were used, significant visual capacity for movement discrimination was demonstrated in five of the animals.

5 These results are discussed in relation to existing theories concerning the function of the cortex in vision.

REFERENCES

1 BARRIS, R W Disposition of fibers of retinal origin in the external geniculate body *Arch Ophth.*, 1936, **14**, 61-70

2 BROUWER, B Anatomical, Phylogenetical, and Clinical Studies of the Central Nervous System. Baltimore: Wilkins, 1927. Pp. 67

3. BYRNE, J G Studies on the Physiology of the Eye London Lewis, 1933. Pp xii+428

4. CAMPBELL, A. W Histological Studies on the Localization of Cerebral Function. Cambridge: Cambridge Univ Press, 1905. Pp. xiv+360

5 Cate, J ten. Akustische und optische Reaktionen der Katzen nach teilweisen und totalen Extirpationen des Neopalliums *Arch néer Physiol*, 1934, **19**, 191-264

6. Dusser de Barenne, J G Corticalization of function and functional localization in the cerebral cortex *Arch Neurol & Psychiat*, 1933, **30**, 884-901

7 Ganser, S Ueber die periphere und centrale Anordnung der Sehnervenfasern und uber das Corpus bigeminum anterius *Arch. f. Psychiat u. Nervenkrank*, 1882, **13**, 341-381

8 Ghiselli, E E The superior colliculus in vision *J Comp Neurol*, 1937, **67**, 451-467

9 Henry, F M Objective indicators and recorders of electric shock *J. Gen. Psychol*, 1935, **13**, 219-222.

10 Henschen, S. E On the visual path and center *Brain*, 1893, **16**, 170-180

11 Holmes, G The cortical localization of vision *Brit Med. J*, 1919, **2**, 193-199

12 Keller, A D, & Stewart, L The superior colliculus and the pupillary light reflex in the cat *Amer J Physiol*, 1932, **101**, 64.

13 Kennedy, J L, & Smith, K U Visual thresholds of real movement in the cat *J Genet Psychol*, 1935, **46**, 470-476

14. Kennedy, J L. An experimental study of the lower absolute thresholds of visual real movement discrimination in the cat Library of Brown University. M. A. Thesis, 1936. Pp. vi+45.

15. ——— The nature and physiological basis of visual real movement discrimination in animals *Psychol Rev*, 1936, **43**, 494-521.

16. Kluver, H An analysis of the effects of the removal of the occipital lobes in monkeys *J. of Psychol*, 1936, **2**, 49-61

17. ——— Certain effects of lesions of the occipital lobes in macaques *J of Psychol*, 1937, **4**, 383-401

18 Koudrin, A N Conditioned reflexes in the dog after extirpation of the posterior halves of the hemispheres. Petrograd Thesis, 1911

19. Lashley, K. S. The mechanism of vision II The influence of cerebral lesions upon the threshold for discrimination for brightness in the rat *J. Genet Psychol*, 1930, **37**, 461-480.

20. ———. The mechanism of vision IV The cerebral areas necessary for pattern vision in the rat. *J Comp Neurol*, 1931, **53**, 419-478

21 ——— The mechanism of vision: XII. Nervous structures concerned in the acquisition and retention of habits based on reactions to light *Comp. Psychol Monog.*, 1935, **11** Pp 79

22 Lashley, K S, & Frank, M The mechanism of vision X Postoperative disturbances of habits based on detail vision in the rat after lesions in the cerebral visual area. *J Comp Psychol.*, 1934, **17**, 355-380

23 Loeb, J. Comparative Physiology of the Brain New York Putnam, 1900. Pp x+309

24. Magoun, H. W Maintenance of the light reflex after destruction of the superior colliculus in the cat *Amer J. Physiol*, 1935, **111**, 91-98.

25 MARQUIS, D. G. Effects of removal of the visual cortex in mammals with observations on the retention of light discrimination in dogs *Proc Asso Res. Nerv & Ment Dis*, 1932, 13, 558-592.

26 ———. Phylogenetic interpretation of the functions of the visual cortex *Arch Neur. & Psychiat*, 1935, 33, 807-815

27 MARQUIS, D G, & HILGARD, E R Conditioned responses to light in dogs after removal of the visual cortex *J Comp Psychol.*, 1936, 22, 157-178.

28. ——— Conditioned responses to light in monkeys after removal of the occipital lobes. *Brain*, 1937, 60, 1-12.

29 MINKOWSKI, M Experimentelle Untersuchungen uber die Beziehungen der Grosshirnrinde und der Netzhaut zu den primaren optischen Zentren, besonders zum Corpus geniculatum externum *Arb aus dem Hirnanat Inst Zurich*, 1913, 7, 259-362.

30 ———. Über den Verlauf, die Endigung und die zentrale Reprasentation von gekreuzten und ungekreuzten Sehnervenfasern bei einigen Saugetieren und beim Menschen *Schweiz. Arch f Neurol. u Psychiat.*, 1920, 6, 201-252.

31. MUNK, H. Ueber die Funktionen der Grosshirnrinde. Berlin Hirschwald, 1881. Pp. iii+131

32. OVERBOSCH, J F A Experimenteel-anatomische onderzoekingen over de projectie der retina in het centrale zenuwstelsel. Amsterdam: Inaug Dissert, H. J Paris, 1927. [Cited by Brouwer (2)]

33 PENARD, S Zur Frage der Beziehungen zwischen Sehrinde und primaren optischen Zentren auf Grund von experimentellen Studien am Affen *Schweiz Arch f. Neurol u Psychiat*, 1935, 36, 131-163

34. POLIAK, S. Die Verbindungen der Area Striata (intrahemisphereale, kommissurale, palliodienzephalische, palliotektale Fasern) bei der Katze und deren funktionelle Bedeutung. *Zsch f. d ges. Neur. u Psychiat.*, 1926, 100, 545-563.

35. ———. An experimental study of the associational, callosal and projectional fibers of the cerebral cortex of the cat. *J. Comp Neur*, 1927, 44, 197-257

36 RANSON, S. W., & MAGOUN, H. W The central path of the pupilloconstrictor reflex in response to light *Arch. Neur & Psychiat*, 1933, 30, 1193-1204.

37. SMITH, K. U. Apparatus for the study of sensory discrimination in mammals. *Science*, 1935, 82, 423-425.

38. ———. The mechanism of discrimination in the cat. Library of Brown University: Ph D. Disser, 1935. Pp 180.

39. ——— Visual discrimination in the cat. IV The visual acuity of the cat in relation to stimulus distance *J Genet Psychol*, 1936, 49, 297-313.

40 ——— The post-operative effects of removal of the striate cortex upon certain unlearned visually controlled reactions in the cat *J Genet Psychol.*, 1937, 50, 137-154.

41 SPENCE, K. W, & FULTON, J F The effects of occipital lobectomy on vision in chimpanzee *Brain*, 1936, 59, 35-50.

42 THIEULIN, G. Recherches sur le globe oculaire et sur la vision du chien et du chat Paris. Danzig, Thèse vétérinaire, 1927 Pp. 93

43 VAN HERK, A W. H, & TEN CATE, J On vision of rabbits after removal of the visual cortex. *Acta Brev. Néer*, 1933, **3**, 96-98.

44 VINCENT, S. B The mammalian eye *J Anim. Behav*, 1912, **2**, 249-255

45. WINKLER, C, & POTTER, A An anatomical guide to experimental researches on the cat's brain Amsterdam Versluys, 1914

Department of Psychology
Stanford University, California

The Journal of Genetic Psychology, 1939, 54, 151-164

FURTHER EXPERIMENTS IN SOCIAL BEHAVIOR AMONG DOGS*

Anatomy-Farm, Cornell University Medical College

W T JAMES

A INTRODUCTORY

The effect of the presence of one dog on another while testing the conditioned salivary reflex was discussed in a previous study (1). It was found that some dogs are inhibited and disturbed by the presence of the second animal. Those most disturbed are the dogs dominated by others in their own group. Those least disturbed are the aggressive and dominant members of the group Between these two extremes are the intermediate type which have little effect on each other These animals give good conditioned reactions and take food without either conflict or submission.

Since publishing the first report we have studied a second group of dogs in an effort to further determine the elements involved in this social behavior It is important to know whether an extremely inhibited individual will become more aggressive when transferred to another group or whether it will remain inhibited in all associations By using two litters of dogs in which the same general social arrangement has developed there is an opportunity to test whether the timid dogs are inhibited because of the aggressive animals of their own group, or whether they are equally disturbed by the strange aggressive individuals from other groups as well. It is also important to know whether there is a variable factor which tends to alter the social standing of a dog. Size might be significant in this regard Another point of investigation was the correlation, if any, between the social position of a dog and its ability to form relatively simple conditioned reactions in the natural environment of the kennel as contrasted with its behavior in the artificial surroundings of the laboratory In the course of the experiments observations on the relation of sex reactions to social dominance were made These concern the effects of aggressiveness on one of the basic behavioral reactions of the organism

---

*Received in the Editorial Office on April 6, 1938

In our study of dogs of different constitutional types at the Cornell Anatomy Station it has been found that these animals fall into distinct behavioral classes. Two pure breeds showing great contrast in behavior are the bassethound and the German shepherd These breeds are also contrasted physically. The bassethound is stocky in body, short legged, and has a slight acromegalic tendency It is also inactive and sluggish in behavior. The German shepherd dog is thin of body, with long legs, and is highly active, showing the characteristics generally associated with alertness. When these two types are crossed the hybrids of the second generation show wide variation both in physical form and behavior. Great differences in social reactions among these hybrids also occur The litter used in the first study were obtained by breeding together two back-crossed basset-shepherd-bassets These five dogs exhibited the differences in social reactions indicated in Figure 1. This diagram shows the relation of each member of the group to all the other members The arrows in the diagram point to the dogs dominated by those from which the arrows arise The top circle represents 1297♂ giving oft four arrows to dominate its four sibs while 1301♀ at the bottom is the most submissive, being pointed by arrows from all the other four.

The six dogs used in the present experiments are hybrids obtained by breeding a three-quarter bassethound-German shepherd with a basset-shepherd F_1. These animals do not differ widely in physical type, but show great variety in social reactions.

Observations were made at intervals throughout the day over a long period of time to find the relationship of each dog to all other members of the litter The most pronounced reactions occur when food is placed in the pans When only a single pan with food is placed in the pen, the aggressive dogs come forward to eat while the timid animals remain at a distance. The aggressive dogs run to meet the kennel man when he approaches while the others are shy and retiring. The animals were later separated into pairs so that each member was observed in relation to every other member of the litter. In this manner it is easy to determine which dog is dominated by all other members of the litter and which is the most aggressive. The relationships between the six members of this litter are indicated in Figure 2. There are no arrows between 1514♀ and 1515♀ since these two animals are apparently indifferent to one another, giving no signs of conflict or aggression Neither will yield when the other

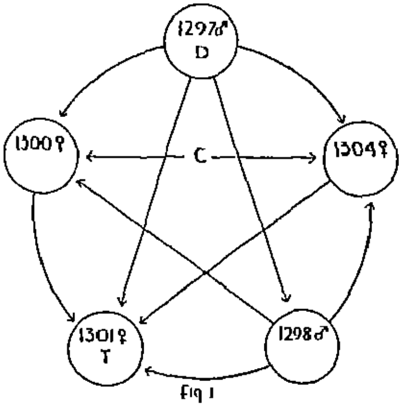


fig 1

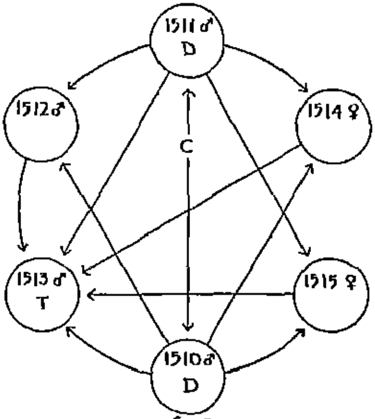


fig 2

FIGURES 1 AND 2

approaches, yet neither one attempts domination over the other. Again one dog in this group, 1513♂ is dominated by every other member of the group just as 1301♀ was in the former group The second group differs from the first by having two equally dominant dogs, 1510♂ and 1511♂, these dominate over the other four members but fight between themselves

After the relationships between the members of this second group had been learned, both groups were brought together in order to find how the aggressive and timid individuals from each would behave toward the strangers. Two animals of Group 2 were placed successively with each dog of the other group until they had been compared with all five of the first group. Thus the three sets of two of the second litter were studied in relation to every member of Group 1, making a total of 15 separate social situations. Each set of three dogs remained together until there was no doubt of their relationships The extremely inhibited type could be detected in a relatively short time However, a week or more was required before the social dominance could be definitely established among some of the more aggressive dogs. The relations in the presence of food were used in these determinations just as in the analysis of the social positions of the members in the two litters. The relative position of each dog among the 11 was finally determined. Table 1 indicates the way in which each of the eleven animals reacted toward the members of the strange group. In the first column the dogs are listed by their recorded numbers, and the following columns show the number of individuals in the strange group to which they were submissive, indifferent, on equality with, or dominated. The dominant dog of each group is marked "*D*," and the most inhibited "*T*."

TABLE 1

Group 1	Submissive	Indifferent	Equality	Dominated
1300♀	1	0	2	3
1301♀T	3	3	0	0
1298♂	2	0	0	4
1304♂	2	2	0	2
1297♂D	0	0	1	5
Group 2				
1512♂	4	1	0	0
1513♂T	4	1	0	0
1510♂D	0	0	1	4
1515♀	3	2	0	0
1514♀	4	0	0	1
1511♂	0	0	2	3

Dog 1297♂, the dominant member of Group 1, dominated five animals of Group 2, and was on equality in conflict with one, 1510♂, which was one of the dominant two of the second group. In the same manner, the most aggressive male of Group 2, this 1510♂, dominated all but one of the dogs of Group 1 The most inhibited and timid dog of each group, namely 1301♀ of Group 1 and 1513♂ of Group 2, were dominated by most of the strangers.

B. CORRELATION BETWEEN SOCIAL DOMINANCE AND THE CONDITIONED FOOD TAKING BEHAVIOR

It is of interest to know if there is any correlation between social position as here expressed and the ability to form conditioned reactions in the natural environment of the kennel The general behavior, including orientation, approach to the food pans, and food taking, was used as an indication of the conditioned reaction rather than the flow of saliva which was employed in the first study. A buzzer was used to signal the food. A special food table was placed in the pen, this consisted of a revolving disk holding six food pans The disk projected from below a partition so that only a single pan of food was presented at a time. The apparatus was operated from a nearby building and the behavior of the dogs could also be watched while the observer was concealed within this building. The observations were made each day at the regular feeding time. The animals were arranged into groups of three For purposes of clearness the most significant observations will be discussed separately.

1. *Behavior when an Aggressive Dog, 1510♂ Was Placed with Two Inhibited Dogs, 1514♀ and 1512♂*

Before starting the experiments the animals were allowed to remain in the run until they had become accustomed to the food table. The first time the conditioning signal was given all three dogs showed an orientation response. Since food had not previously been given in this apparatus, the dogs which took food were only those close enough to see the pan rotate into position. On the first occasion it happened that the two less aggressive dogs, 1514♀ and 1512♂, were standing near the pan Both of them saw the food and began to eat from the single pan During the first two presentations of food the aggressive animal, 1510♂, did not eat but he did investigate the pan, and on the following six feedings he alone ate all of the food. His approach to the pan began when the signal sounded and as he approached the submissive animals immediately withdrew. There was no physical conflict of any kind.

All three dogs formed the conditioned approach reaction on the first day after nine signals. In every instance the buzzer elicited the orientation and approach reaction, although following the first two presentations of food the eating reaction was accomplished only by the aggressive dog This general behavior continued for four consecutive days.

2 *The Two Inhibited Animals, 1514♀ and 1512♂, in Association with the Extremely Inhibited Animal, 1513♂*

The object was to find, first, how the two dogs, 1514♀ and 1512♂, would react after the aggressive dog, 1510♂, was removed and, second, to determine whether the extremely timid animal, 1513♂ of Group 2, would take food in company with other shy animals. Both 1514♀ and 1512♂, having already formed the conditioned reaction, ran to the food table as soon as the buzzer sounded. The timid dog never approached the pan In this manner 1513♂ was deprived of food for five days, after which time 1514♀ and 1512♂ were removed. For seven days, when 1513♂ was alone in the run, food was placed on the table in full view so that the dog could see the procedure, but even then the dog would not take food while the experimenter was present. When the experimenter withdrew the dog slowly approached the pan and ate hesitantly The animal was timid and would retreat when any sound occurred Every effort to condition the reaction to the sound of the buzzer failed Under normal kennel conditions as well as under laboratory conditions, as found in the previous report, dogs of this disposition are handicapped and limited in their behavioral reactions.

The aggressive dogs of the group take the lead in all activities and respond first to all new situations. Aggressiveness is usually associated with alertness, and as a rule the aggressive animals are more capable of learning In some cases, however, the dominated dogs may be as capable of learning as the aggressive ones, but they are always handicapped in the presence of the aggressive animals The extremely inhibited animals are not only disturbed by the presence of other dogs but also by people and any variation in their environment.

We have attempted to determine the physical characters which may effect the social positions and general behavior of dogs within a group The relative size of an animal may be important in this connection

C THE RELATIONSHIP OF SIZE TO SOCIAL DOMINANCE

The weights of each dog in groups 1 and 2 are given in Table 2. In Group 1 the dominant male weighs less than any of the other dogs except 1301♀, the timid female. Since the other two females are about equal in weight and the dominant male weighs less than the second male of the group, it cannot be concluded that size determines the social reaction. In Group 2, 1510♂ and 1511♂ dominated all other members of the group and fought each other. These were the heaviest dogs of the group. The timid male of Group 2 weighs more than two others of the group although this dog is dominated by these two. On the basis of weight alone this animal should have maintained a better social position.

TABLE 2

Animal No	Kilos	Animal No	Kilos
1301♀T	17	1515♀	17.2
1304♀	26.8	1514♀	18
1300♀	27.8	1510♂D	23.6
1298♂	21.3	1513♂T	19.3
1297♂D	19.5	1512♂	19.3
		1511♂	23.8

When the weights of the dogs of the two groups are compared it is found that the dominant male, 1297♂, weighs less than three dogs and practically the same as two others, yet all of these are dominated by him. Comparing the social reactions and weights of this group of 11 animals would seem to indicate that size plays little part in dominance. This conclusion cannot be applied to all dogs however until animals which show greater difference in weight have been considered. For this reason some of the aggressive animals were pitted against much larger dogs from different litters, and the timid were compared with smaller aggressive dogs.

1 *The Dominant Dog, 1297♂, Competing with a Larger Hybrid Dane-St Bernard, 1612♂*

The dominant dog of Group 1, 1297♂, weighing 19.5 kilos, was placed in the run with a Great Dane-St Bernard, No 1612♂, which weighed 60 kilos. The latter animal was not extremely aggressive within his own group, although he was not dominated by any of them. The two dogs were placed together before and during feeding time for a number of days. At first they apparently paid little attention to each other. A cracker thrown into the run was quickly eaten by 1297♂. A single pan of food was then placed in the run and again animal 1297♂ began eating, at the same time growling at 1612♂. The larger dog withdrew. After this, food was placed in a long trough from which both dogs could eat, but 1297♂ took most of the food At another time a pan of food was placed near the larger dog so that he might begin eating before 1297♂ could reach the pan. He began eating but 1297♂ approached and tried to push him away For a brief interval both dogs growled with snarled lips and raised hair, the typical attack attitude. A fight began and soon the larger dog stood over the smaller animal holding him down, biting him viciously around the shoulders and neck Being unable to free himself from the larger dog, he was badly beaten. Realizing his disadvantage, 1297♂ ceased to growl and 1612♂ released him Following this conflict the larger animal now ate all the food while 1297♂ retired and looked on. The food was presented in the same way on the following day and the larger dog ate the food without protest from the smaller animal

Precisely the same behavior occurred when the second aggressive dog, 1510♂, was placed in the run with 1612♂. There was this difference however, on first meeting, animal 1510♂ stood on his hind feet with front feet on the shoulders of 1612♂ and licked him around the face and neck. This would indicate that the size of the larger dog did not intimidate the aggressive animal After the first conflict over the pan of food, in which the smaller animal fared similarly to 1297♂ above, he then shied when the larger dog growled and approached.

2. *Two Unaggressive Dogs 1512♂ and 1514♀, placed with the Larger Dog, 1612♂*

When the unaggressive dogs, 1512♂ and 1514♀, were placed with the larger dog, 1612♂, they behaved as they did towards the unaggressive members of their own group. They would withdraw if 1612♂ protested, but would eat with him if he permitted.

3. *A Small Aggressive Animal in Company with a Large Unaggressive Dog*

If size were the important factor, the large timid dog should not withdraw from food when the small aggressive dog protests To test this, the two timid dogs, 1513♂ and 1512♂, from Group 2 were placed in the run with a hybrid Pomeranian-Pekingese F_1, 1326♂, which weighed 7 2 kilos This small dog was so aggressive among the members of his own group that it was necessary to separate him from them. When placed in the run with the two larger dogs weighing 19.3 kilos each, the small dog took food from the pan or trough without hesitation and would prevent the two larger animals from approaching the food by assuming an attack attitude (Figure 3)

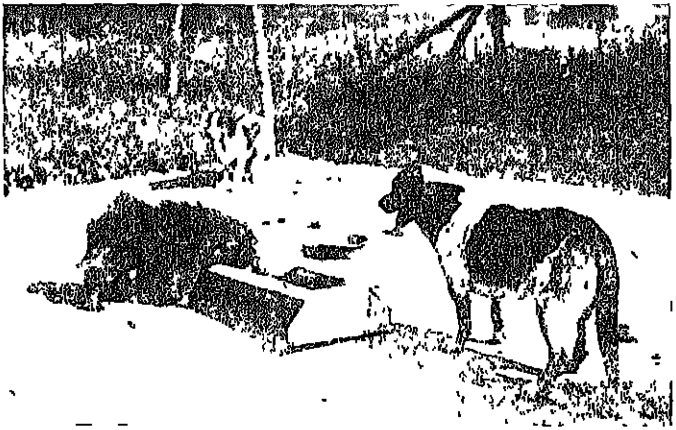


FIGURE 3
THE SMALL AGGRESSIVE DOG EATS FROM THE TROUGH WHILE THE LARGER BUT UNAGGRESSIVE ANIMALS REMAIN AT A DISTANCE

After comparing the small animal with the larger timid dogs, the extremely aggressive dog 1510♂, was placed in the pen with the three. There was conflict over the food between the aggressive dogs and during the fight the smaller dog was held down by the larger one and severely bitten. Before the two could be separated 1326♂ was so badly hurt that the experiment was discontinued. Enough evidence had been obtained to show that the aggressive reaction, involving a complicated behavioral system was, by far, more important than size in social conflict. However, it cannot be said that size and fighting ability have no part in social dominance since aggressiveness may be unconditioned by repeated failure in combat

There is also the question whether the attack reaction of the small dog towards a larger dog which is inhibited after an unsuccessful engagement refers to all larger animals or whether the inhibition is expressed only toward the victorious animal With this in mind, Nos 1297♂ and 1510♂, the dogs which had been beaten by 1612♂ and no longer challenged this larger animal, were placed in a run with another Great Dane-St. Bernard, 1921♂, weighing 45 kilos. The latter dog was one of the most aggressive animals in his group The behavior of 1297♂ and 1510♂ toward 1921♂ proceeded along exactly the same course as had been followed with 1612♂. Both 1297♂ and 1510♂ first assumed the aggressive attitude, but again found that they were at a disadvantage in conflict and lost their aggressiveness toward the larger dog. The attack attitude toward the second large animal was inhibited by specific encounter and was not brought about simply by the former experience with the other large dog This would seem to indicate that the unconditioning is specific toward the given victor rather than general for all larger dogs.

D Social Dominance and the Mating Reaction

Numerous investigators have employed the sex drive in learning experiments as well as social behavior. During the course of the above studies observations were made on the relation between aggressiveness and sex behavior A female of this group when in oestrus was placed in a run with three males: an extremely timid male, 1513♂; a less timid male, 1512♂, and a dominant male, 1510♂ It had previously been found that the timid male, 1513♂, had a vigorous mating reaction However, in the presence of the aggressive dog, 1510♂, this timid animal made no attempt to mate. Both timid dogs stood aside while 1510♂ mated with the bitch. The nearness of approach on the part of each of the three animals was significant. The timid male remained in the far corner of the run while the less timid dog came closer Occasionally 1512♂ would approach and try to mount the female, but in every case the aggressive dog growled and 1512♂ assumed the passive defense reaction (Figure 4) The aggressive dog was now removed in order to observe the reactions of the two timid animals Within five minutes after this dog was removed, the timid animal, 1513♂, mated with the bitch During this time animal 1512♂ remained near and at

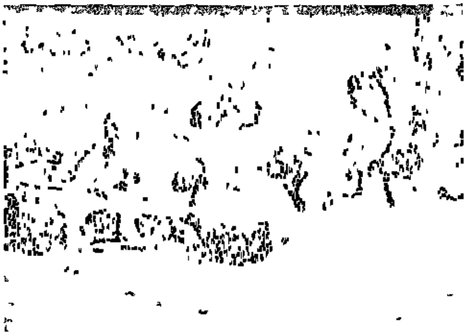


FIGURE 4
THE DOMINATED DOG, WHO WAS TRYING TO MATE WITH THE FEMALE IN THE BACKGROUND, DROPS DOWN IN A PASSIVE DEFENSE ATTITUDE AS THE AGGRESSIVE DOG APPROACHES

intervals would also try to mount the female Later, on the same day, 1512♂ mated with the female in the presence of 1513♂ There was no combative reaction between these dogs. Since the extremely inhibited animal was disturbed by the presence of persons on other occasions, a kennel man was allowed to enter the run during one of the matings Animal 1513♂ immediately left the female and retired to a distant corner of the run.

It was conclusive from the above social situation that the sex behavior, one of the strongest reactions of the dog, is completely inhibited in a timid animal by the presence of a socially dominant dog This reaction, like others, is also disturbed by change in the environment

E Discussion and Conclusions

In these experiments we have made a study of the social behavior of dogs under kennel conditions rather than in the laboratory as in the previous study Two groups of animals were used in which a similar social order had developed In each group the extremely aggressive dogs dominated all the other members. There were also extremely timid dogs which were dominated by every other animal The individuals of the intermediate type formed a submissive-dominance order among themselves.

When the members of the first group were compared with those of the second group, the aggressive dogs had conflict between themselves and dominated the less aggressive strangers. If the aggressive dogs were removed, one of the less aggressive members assumed the dominating position. This shift in social position occurred only among the dogs of the mid-group The most submissive animals maintained the same relative position among strangers

It was found that size was not of first importance as a determinant of social standing. A small aggressive animal dominated larger animals that were less aggressive than himself. In some cases this aggressive reaction toward the larger dog was unconditioned after failure in conflict, but the behavior was inhibited only toward the particular animal and was not generalized to all larger animals

There seems to be some inherent nervous factor or group of factors which determines social position. Evidence for this is found particularly in the case of the submissive animals who do not try to dominate each other even when they are isolated. These dogs remain inhibited and unassertive in any group. They are not only disturbed by the presence of the aggressive animals but are likewise disturbed by timid dogs from other groups It might be thought that they have lost their aggressive tendency as a result of past experience with the more aggressive members of their own litter. This may be so, but we have found no evidence for generalization either of positive or negative social reactions. Aggressive behavior may be unconditioned toward a second individual as shown in the experiment on size, but the reaction does not become inhibited toward other dogs thereafter The timid dogs are disturbed by every change in their environment even when alone. They hesitate to eat from a new food pan, or from the regular food pan if it is placed in a different position. It was shown that they do not form a specific food taking reaction to signals. They will not eat food from the pan until they learn by repeated approaches and withdrawals that there is no dangerous object nearby. This behavior occurs every day. Even while eating the dogs withdraw immediately at any sounds Every experiment involving these submissive animals indicates that they are an inferior type. This has been found both in the food taking behavior and in the mating reaction.

This study emphasizes, as did the previous one, the significance of constitutional differences in social behavior Experiments dealing with the stimulus situation, considering conditioning as the sole basis of behavior, does not adequately explain the above behavior differences The specific conditioned reactions and social behaviors are manifestations of definite neuro-physiological systems peculiar to each type It is true that conditioning plays a part in social reactions, but it does not fully interpret the interaction between the dogs. Nor can it be said that the difference is simply one of "drive," or of inhibition excitation ratio. There are specific quantitative and qualitative aspects of each organism which places it in its social position For these reasons an understanding of the submissive and retiring type and of the aggressive and domineering type may possibly be reached by a consideration of the genetic factor in addition to the stimulus or environmental factor. Further experiments will deal with the possible influences of glands of internal secretion on the several types isolated in these social experiments

F. SUMMARY

1. Similar social orders were formed in a litter group of five dogs and in a second group of six dogs
2. In this social order the most aggressive animals dominated all members of the group. The extremely timid and inhibited animals were dominated by every member.
3 The extremely aggressive dogs remained aggressive and the submissive type remained inhibited among strangers.
4. There were intermediate types which shift their social position as they are transferred from one group to another.
5. In the kennel, as in the laboratory, the aggressive dogs always have the advantage in any conditioned action Animals of the intermediate type form the conditioned reaction to the signal but it is inhibited by the presence of the aggressive dogs.
6. Size was considered as a possible factor which might effect social position.
7 An aggressive animal challenged a larger animal over food. The small dog lost his aggressiveness towards the larger animal after repeated failure in conflict
8 This aggressiveness was unconditioned only towards the particular animal involved and was not transferred to other large animals
9. The extremely timid dogs never attempted to dominate a smaller aggressive animal

10. In the mating reaction, as in the food taking performance, the dominated dog was at a disadvantage. The inhibited male would mate with the female when alone with her, but the reaction was inhibited in the presence of those animals which dominate him. The animals of the intermediate type mate more freely, but when an extremely aggressive animal approaches they also become inhibited

11. This study emphasized, as did the previous one, the significance of constitutional differences for behavioral observations.

REFERENCE

1 JAMES, W. T The effect of the presence of a second individual on the conditioned salivary response in dogs of different constitutional types J Genet Psychol, 1936, 49, 437-449

Anatomy-Farm
Cornell University Medical College
Mohegan Lake, New York

The Journal of Genetic Psychology, 1939, 54, 165-180

HEIGHT-WEIGHT VARIABILITY FROM BIRTH TO FIVE YEARS OF AGE FOR CHILDREN REARED WITH OPTIMAL PEDIATRIC AND HOME CARE*

College of the City of New York, and Port Chester, New York

JOHN GRAY PEATMAN AND REGINALD A. HIGGONS, M D.

A. INTRODUCTION

In a previous study (7) of growth norms from birth to five years of age for children reared with a relatively optimal degree of pediatric and home care, at least in factors of diet, sunshine, exercise, and rest, we obtained results which contrasted significantly with the customarily accepted norms of height-weight growth at these ages. A comparison of our obtained average heights and weights for boys and girls with the results of other investigators revealed that our groups, after three months of age, were increasingly above the average heights and weights for given ages up through five years This contrast is summarized by Figure 1. Woodbury's norms (8) are used for the comparisons. And his results are very similar to those of Baldwin and Wood (3) (except that after the third year Baldwin and Wood's height curves tend to be about 1 inch lower than Woodbury's) and to those, for height, of the New York State Department of Health (5), whose norms were developed from data of Holt, Woodbury, Emerson and Manny, Crum, Smith, Baker, and Lucas

In the present article we shall present and analyze the *variability* of our height and weight measures, comparing our results with those of Woodbury (8). Since average heights and weights change with age, we shall consider both so-called *absolute variability* [in terms of the standard deviation (σ) from the mean] and *relative variability* (in terms of the per cent ratio of σ to its mean).

The nature of our population sample, the processes of measurement, and type of pediatric and home care have been described in our aforementioned article (7) and will not therefore be repeated in detail here. Briefly, our sample consisted of 5149 height-weight pairs of measurements, taken of 1112 boys and girls, 1 to 60 months

---

*Received in the Editorial Office on April 25, 1938

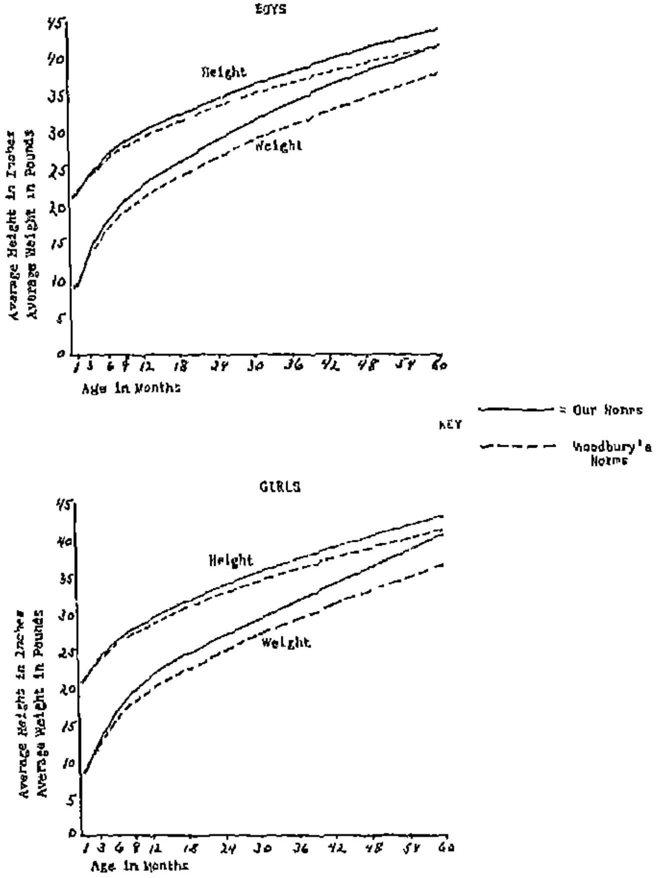


FIGURE 1

COMPARISON OF OUR HEIGHT AND WEIGHT NORMS WITH WOODBURY'S

of age, growing up in an area suburban to New York City. All of the subjects were patients in the private practice of the pediatrician and were, on the whole, of homes of at least middle class standards of living. Twenty-two age groups were obtained and differentiated as to sex—the 44 groups averaging 117 cases each. For the first 12 months, age groups at each month included cases measured within two weeks of the particular month birthday (the tendency in pediatric practice is to examine infants nearer the month birthday than around the middle of the month birthday). The remaining 10 age groups, viz., 15, 18, 21, 24, 30, 36, 42, 48, 54, and 60 months, included cases measured within one month of the particular age.

In the following pages we shall consider the absolute and relative variability of our groups in respect to the following questions: (*a*) Does their variability in height and weight tend to increase, decrease, remain constant, or vary in any other manner with increase in age? (*b*) Are there any sex differences in the trend of variability with age? (*c*) Do our groups, subjected to a relatively optimal degree of pediatric and home care, tend to vary with age in a manner that contrasts with or is similar to the trends of variability obtained by Woodbury for children reared under more ordinary growth conditions? And, finally, we shall consider our results and Woodbury's in relation to the question of the possible effects of relatively optimal pediatric and home care on variability in height and weight during the first five years of childhood.

B. Results

The results of our analysis are presented graphically in Figures 2-5. We have used graphs rather than tables to portray the trends of our results, since trends are more obvious by diagrams than by tabulation. The numerical values of σ and V can of course be read from the graphs.

Our height and weight distributions, for both sexes and for all age groups, were analyzed for their possible divergence from the theoretical *normal* probability distribution. In all cases the distributions are not significantly skewed (i.e., the ratios of Sk/σ_{Sk} are less than 3.0), and the distributions do not deviate significantly from mesokurtosis [i.e., the ratios of $(\text{Ku}-263)/\sigma_{Ku}$ are less than 3.0].

1. *Interpretation of Figure 2—Variability in Height*

Our results for both absolute and relative variability in height of boys and girls, 1 to 60 months of age, are presented in Figure 2. An analysis of the trends and differences indicates:

1. There are no consistent sex differences in height, either for absolute or relative variability. However, there is a *tendency* towards slightly greater relative variability in the girls' heights.
2. For both sexes there is a definite trend from 1 to 60 months

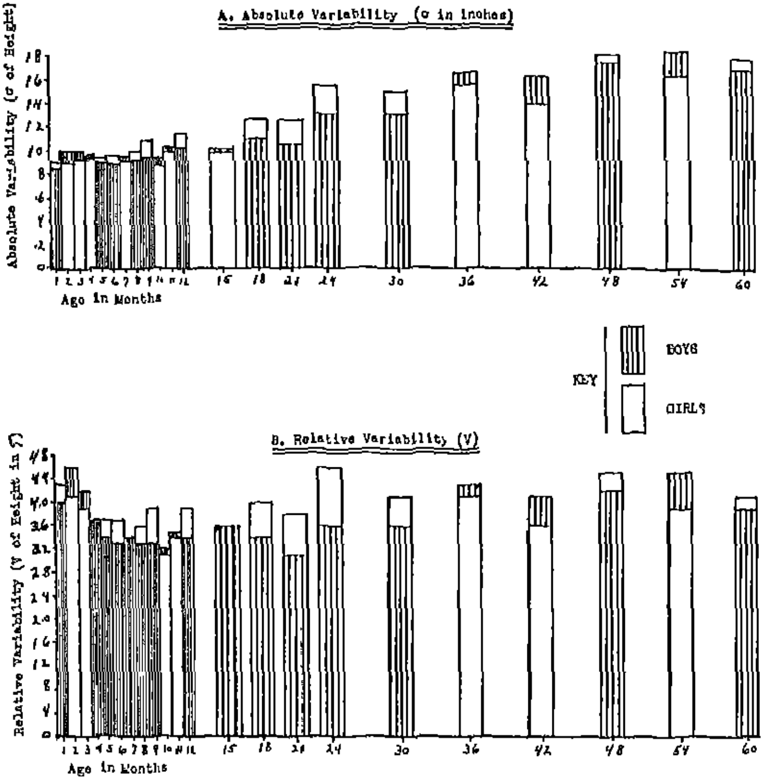


FIGURE 2

Trends of Variability in Height for Boys and Girls from 1 to 60 Months of Age

of age toward increasing amount of absolute variability. This trend runs from an average of about .9 inch for the first six age groups (1 to 6 months) to about 1 and ¾ inch for the last two age groups (54 and 60 months). (Cf. also averages of our data as presented in Figure 4.)

3. For both sexes, per cent of relative variability begins to decrease during the first several months of age, then there is an irregular but definite increase such that the variability during the later age groups tends to fluctuate about the beginning level.

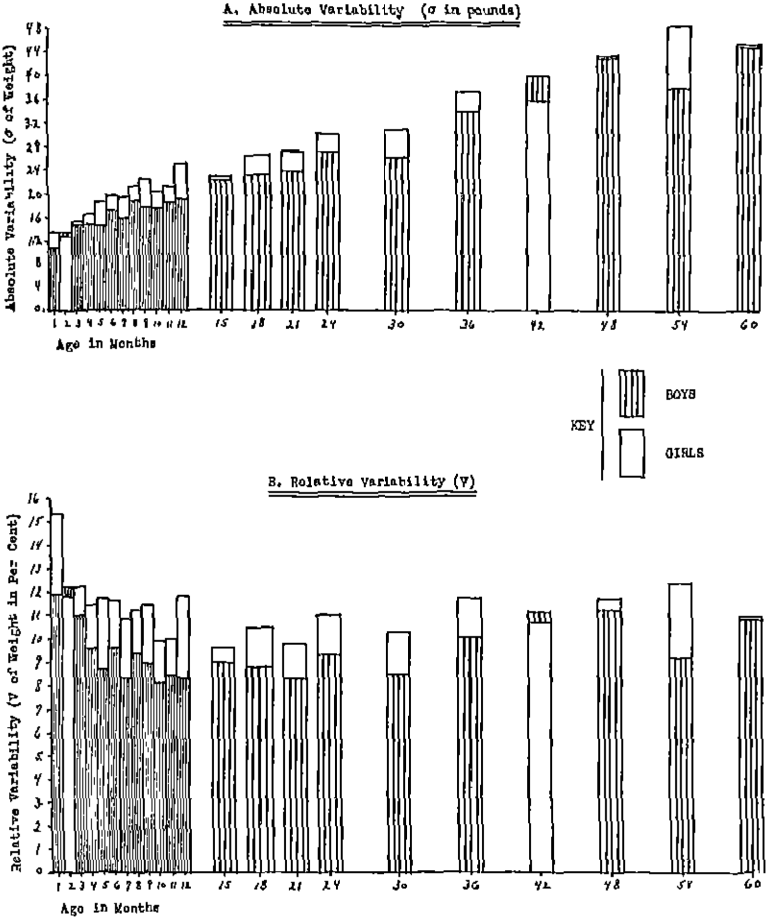


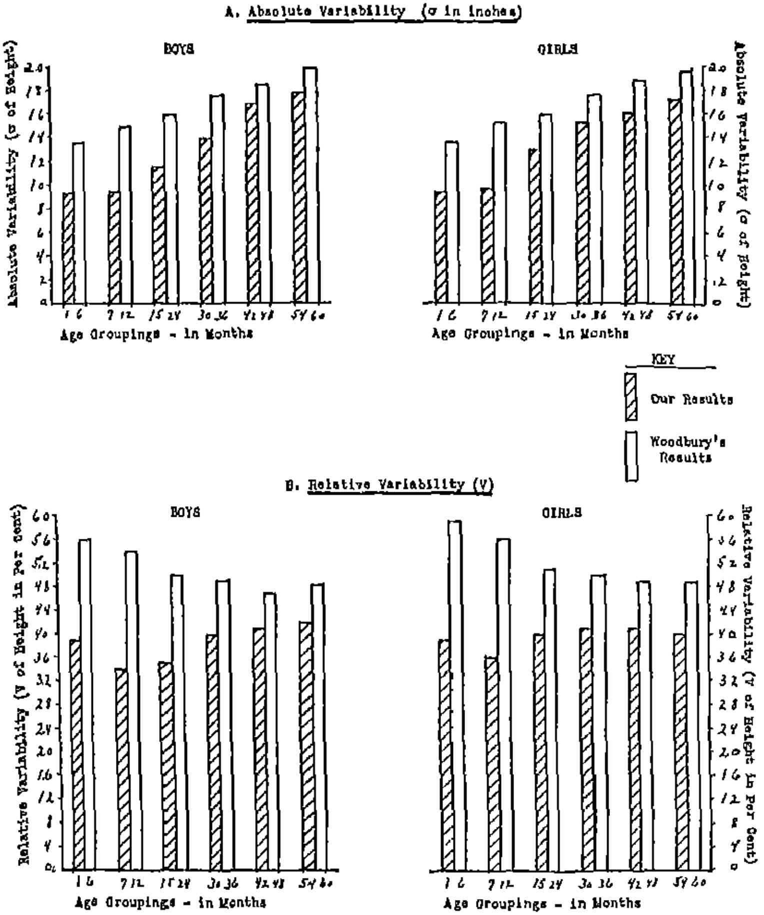
FIGURE 3

TRENDS OF VARIABILITY IN WEIGHT FOR BOYS AND GIRLS FROM 1 TO 60 MONTHS OF AGE

2. *Interpretation of Figure 3—Variability in Weight*

Our results for both absolute and relative variability in weight of boys and girls, 1 to 60 months of age, are presented in Figure 3

1. The girls tend to be consistently more variable at each age


A. Absolute Variability (σ in inches)

B. Relative Variability (V)

FIGURE 4

Comparative Trends of Variability in Height Our Results and Woodbury's

level than the boys This is true of both absolute and relative variability The two age groups, viz., 2 months and 42 months, which are exceptions to this general trend, can be interpreted as insignificant reversals.

2 For both sexes there is a definite and fairly consistent trend

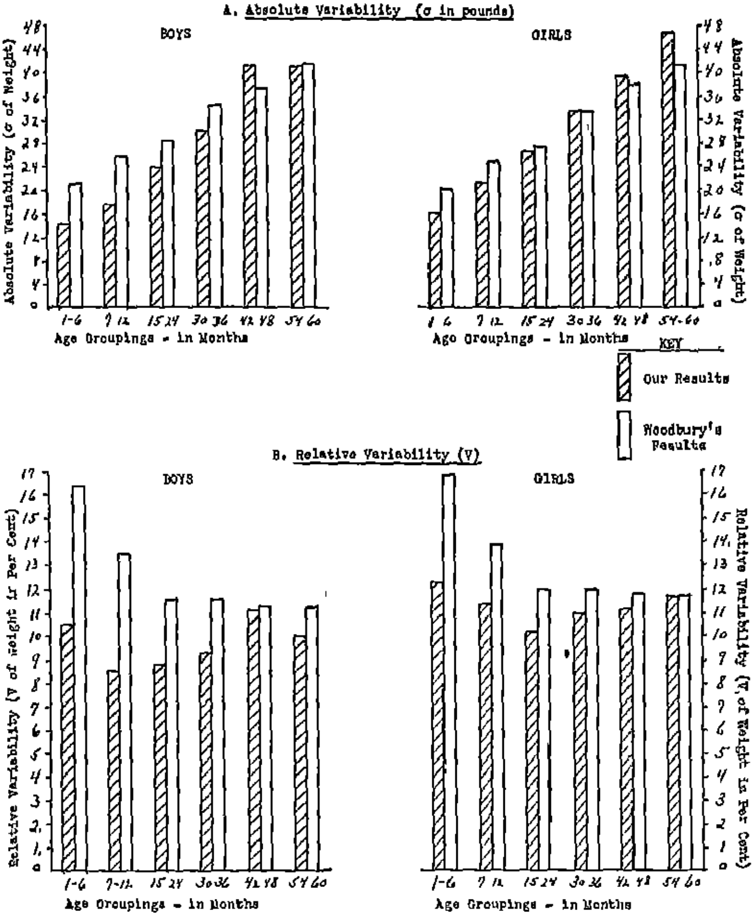


FIGURE 5
COMPARATIVE TRENDS OF VARIABILITY IN WEIGHT OUR RESULTS AND WOODBURY'S

over the 60 month period of growth toward increased absolute variability in weight In the case of the boys, this increase is from about 1.4 lbs for the first six months to about 4 lbs for the last two age groups (54 and 60 months). (Cf also averages of our data as presented in Figure 5.) Similarly, for the girls, the increase is from about 1.6 lbs. in the early months to over 4½ lbs. in the later months.

3. As for relative variability, the results for weight reveal a trend somewhat analogous to that obtained for height in Figure 2. That is, for both sexes the per cent of variability is greater for the beginning age groups than succeeding ones up to about two years. After this decrease during the early period, there is then an irregular but definite increase towards the higher percentage of the beginning level.

3. *Interpretation of Figure 4—Comparisons of Variability in Height*

Our results of variability in height, for both boys and girls, are compared with those of Woodbury in Figure 4.

a. Absolute variability.

(1) For both sexes, Woodbury's data reveal greater amounts of absolute variability in height for each of the six combined age groups.

(2) On the other hand, Woodbury's results, as well as ours, are indicative of a trend of absolute variability which increases with age. For both sexes, the standard deviations average about 1 1/3 inches for the first six months and increase to practically 2 inches for the last six months; whereas our averages go from about .9 inches to 1¾ inches.

(3) Possibly significant, because of its consistency for both sex groups, is the fact that the rate of increase in absolute variability with age is greater for our sample than for Woodbury's. The amount of variability, as we just saw, is nearly doubled for our groups, whereas for Woodbury's groups it is only about half as much greater at the end as at the beginning age level.

(4). Comparing Woodbury's results for the two sexes, we see that there is no consistent difference in the direction of either sex, and all differences are slight. As has already been pointed out, results of a similar kind were obtained in the comparison of our sex groups.

b. Relative variability in height.

(1). Woodbury's results definitely indicate for both sexes a trend in the direction of decreased variability with increase in age. This contrasts with the trend of our results. Although during the first year the variability of our groups begins to decrease in a manner comparable to the decrease of Woodbury's groups, the trend then reverses so that the later age groups are as variable as the first.

(2). For both sexes, Woodbury's age groupings are all relatively more variable in height than our groupings However, with increase in age, there is a definite tendency for the amount of the difference to decrease.

(3) Comparing Woodbury's results for the two sexes, we see that the girls at each age grouping are consistently more variable than the boys, although the differences are slight Although not as consistent, a tendency of greater relative variability for girls' heights is manifest by our groups

4 *Interpretation of Figure 5—Comparisons of Variability in Weight*

Our results of variability in weight, for both boys and girls, are compared with those of Woodbury in Figure 5

a Absolute variability

(1). For both sexes, Woodbury's and our results reveal a trend of increased absolute variability in weight with increase in age.

(2). For both sexes, Woodbury's groups are not consistently more variable than ours His first group (1 to 6 months) is more variable than ours by about 2/5 of a pound This difference practically disappears in the case of the girls at the 15-to-24 months group, and our later groups are somewhat more variable For the boys, Woodbury's groups do not tend to be more variable after three years of age Thus, as was true of absolute variability in height, the rate of increase in absolute variability with age is greater for our groups than for Woodbury's In fact, for weight, our groups tend to be more variable during the fourth and fifth years

(3). Comparing Woodbury's results for the two sexes, we see that at each age grouping, except that of 42-48 months, the boys are slightly more variable than the girls This contrasts with our results In both Figures 3 and 5, our data reveal a consistently greater amount of variability in weight for the girls than for the boys at each age grouping (two minor exceptions revealed in Figure 3 were previously discounted)

b Relative variability in weight

(1). Woodbury's results for both sexes reveal a rapid decrease in relative variability during the first two years, after which time there is very little change. During the first year for boys and during the first two years for girls, the relative variability of our groups decreases in a manner somewhat comparable to the decrease of Woodbury's groups, but thereafter the variability of our groups tends to increase and reveal the per cent of variability characteristic of the first age grouping (1 to 6 months) We saw that analogous contrasting trends were also obtained for relative variability in height

(2). For both sexes, Woodbury's groups are considerably more variable during the first year. Thereafter, the difference decreases and tends to disappear with the later age groups

(3). Comparing Woodbury's results for the two sexes, we see that the girls are slightly more variable at each age grouping. This is in accord with the trend revealed by our results, as previously noted in Figure 3.

C. Summary and Discussion of Results

On the basis of our results and Woodbury's, it is evident that no simple statement will adequately summarize the general nature of variability in height and weight of boys and girls during the first five years of growth. We have seen that in several important respects our groups of children, subjected as they were to a relatively optimal degree of pediatric care and home supervision in conditions of height and weight growth, reveal consistently contrasting differences in variability when compared with the groups of Woodbury's populations, which developed under circumstances certainly more ordinary in respect to conditions of height and weight growth.

In a previous study (7) on growth norms, we compared the average heights and weights of our population with the norms obtained by Woodbury and other investigators. We found that for the first three months of age our groups were not heavier or taller, on the average However, after three months of age, our groups were increasingly above the average heights and weights of the usual norms (*Cf* Figure 1). After a careful analysis of our population, we advanced the hypothesis that environmental conditions arising out of relatively optimal pediatric and home care might account for the increased trends of our height and weight averages Whether, then, the same hypothesis is the best one to account for the contrasts in variability between Woodbury's and our populations, as revealed in this study, is problematical. With this problem in mind, we shall proceed to summarize and discuss our results.

1. *Absolute Variability of Height and Weight*, for both boys and girls, *increases with age* during the first five years of growth. This is true of both our populations and Woodbury's. It seems reasonable that the different conditions of growth for the two population samples would not make for any difference in this respect

2 *The amount of increase in absolute* variability during the five year period is greater for our groups than for Woodbury's Woodbury's groups are, however, consistently more variable in height at each age grouping and more variable in weight during the first two years.

The difference in absolute variability for the earliest age grouping is a statistical statement of the greater heterogeneity in height and weight of Woodbury's groups Although the *averages* in height and weight for the first several months of growth are practically the same (*cf.* Figure 1), the greater variability of Woodbury's groups at this period is to be expected because of the greater number and more heterogeneous character of his measurements [1] The important question here is. Can the greater amount of increase in absolute variability for our populations be reasonably expected to result from the superior pediatric and home care they received?

In dealing with this question, we would note first that the greater variability in weight of several of our later age groupings does not mean that the children at the lower end of their groups weigh less than the children at the lower end of Woodbury's corresponding groups. Greater variability here does not mean this, since the averages of our groups are considerably higher than those of Woodbury's. Thus, comparing directly the weights of boys at 42 months of age, we have the results shown in Table 1 [2]

Here it is evident, assuming a *normal* distribution of cases in each

---

[1] Woodbury's populations were assembled from the records of 172,000 children, being the number of cases satisfying his selective criteria from a total of more than 2 million records sent him at Washington during Children's Year (1918) About 69 per cent had both parents born in the United States The Southern States were under-represented, one-fifth of his cases residing in Iowa and the bulk of the remainder in California and the East-North-Central States

[2] Obtained, rather than smoothed, averages are used for these comparisons The two age groups of Woodbury's data are presented since his averages were assembled in respect to middle-of-the-month birthday. Our data were obtained in reference to the month birthday This difference in points of reference has been taken into account in the comparisons of Figures 4 and 5, as well as of Figure 1

TABLE 1
WEIGHTS OF BOYS AT 42 MONTHS

	Age	Mean	σ	V	Mean±1σ	Mean±3σ
Our Group	42 mo	35.9	4.0	11.1	31.9 to 39.9	23.9 to 47.9
Woodbury's	41½	32.5	3.6	11.1	28.9 to 36.1	21.7 to 43.3
	42½	32.8	3.8	11.6	29.0 to 36.6	21.4 to 44.2

group compared, that even though our group is absolutely more variable, about 78 per cent of our cases have weights above the means of Woodbury's groups (in terms of actual count, 76 per cent of our cases in this group were heavier than 32.75 pounds, and the lowest obtained weight in the group was 27 pounds).

Comparing the weights of girls at 60 months of age, we have the results shown in Table 2.

TABLE 2
WEIGHTS OF GIRLS AT 60 MONTHS

	Age	Mean	σ	V	Mean±1σ	Mean±3σ
Our Group	60 mo	41.4	4.5	10.9	36.9 to 45.9	27.9 to 54.9
Woodbury's	59½	36.3	4.2	11.6	32.1 to 40.5	23.7 to 48.9
	60½	36.8	4.2	11.4	32.6 to 41.0	24.2 to 49.4

Again, on the assumption of *normal* variability, about 85 per cent of our cases have weights above the means of Woodbury's groups (in terms of actual count 84 per cent of our group were heavier than 36.75 pounds; and the lowest obtained weight in the group was 32 pounds).

Important in our consideration of the different amounts of increase in absolute variability is the fact that, *in comparison with Woodbury's groups*, the increase in weight for our groups is relatively greater than the increase in height. This difference is in accord with reasonable expectancy, since factors of diet, sunshine, exercise, and rest, ordinarily have a more immediate and observably greater effect on weight than on height. Furthermore, the effects of factors of diet, etc., on the heights and weights of children are known to be non-uniform. The effects of such factors will be greater for some children than for others, and will probably vary for the same child at different times. And it is reasonable to suppose that, after an initial lag, *superior conditions* of diet, sunshine, exercise, and rest, will make for increases in height and weight which are variable and at the same time greater than those of ordinary expectancy, during at least the first five years of growth. In other words, the fact that the amount of increase in absolute variability over the five year period is greater for our groups than for Woodbury's is a result compatible with the implications of the hypothesis we previously advanced to account for the obtained increases over ordinary expectancy in *average* heights and weights (portrayed in Figure 1)

3 *Relative variability* in height and weight, for both boys and girls, decreases with age during the first year of growth. This is true of both our populations and Woodbury's However, after the first year or so, the relative variability of our groups *increases* to that of the beginning months, tending to stabilize at about that level; whereas, Woodbury's groups continue to *decrease* in relative variability during the second year, with little change after that time during the remainder of the five year period Furthermore, for height, Woodbury's groups are more variable throughout the five year period than our groups However, the difference decreases with age On the other hand, for weight, our groups tend to be as variable as Woodbury's at the end of the five year period, in spite of the fact that his groups are considerably more variable than ours during the first year.

In considering these contrasting trends, it is important to bear in mind that the average heights and weights of our groups are larger than Woodbury's after the first few months. And, since for height, Woodbury's groups are *absolutely* more variable than ours, it is to be expected that his groups will be *relatively* more variable The significant contrast, for both height and weight, lies not so much in the per cent differences between respective age groupings, as in the trends that the results take over the five year period The downward trend of relative variability is not only checked for our groups, but is reversed Whether this reversal can be attributed to the effects of superior conditions of diet, etc., is problematical However, such a reversal is not contradictory to the implications of such an hypothesis

It is possible that the per cent of relative variability in height and weight, for our groups as well as for Woodbury's is levelling out at the 4th or 5th year of age, not to change much for the next several years of growth However, on the basis of studies made nearly half a century ago, Karl Pearson (6) made the statement, "so far as weight and height are concerned female and male both lose variation as they grow older" Such a trend beyond five years of age is more predictable from Woodbury's data than from ours It is of course possible that at later ages such a decrease in trend will occur for our groups. However, an analysis of Garside's tables (4) also reveals no decrease in the relative variability of both height and weight for boys and girls from 36 to 72 months of age Their measures are reported at 3-month intervals for these three years of growth. Relative variability in *weight*, as a matter of fact, for both the boys and girls of these groups, definitely *increases* during this period Garside's subjects are described as in good health, the majority being from "the rather favoring social environment provided by parents of professional and merchant classes"

Baldwin reports results for relative variability in height and weight for 120 children, 7 to 17 years of age, with repeated annual measurements made over a period of from 6 to 10 years on each child (2, Table 26, p 147). For these ten years of growth, relative variability of the girls, for both height and weight, tends to *increase* up through 12 years of age, and then to decrease. For the boys, there is not such a definite trend until after 14 years of age, at which time there is apparently a definite decrease in relative variability for both height and weight

In view of the results of these various studies, Pearson's statement about a decreasing trend of relative variability in height and weight may be definitely characteristic of development following the onset of pubescence, regardless of variations in growth conditions after that time Prior to this period, the trend character possibly varies with different growth conditions,—superior environmental factors of growth tending to make for a trend of increasing relative variability

4 *Sex differences* In *height*, neither sex is consistently more variable, *absolutely*, than the other. However, there is a *tendency* for the girls to be slightly more variable, *relatively*, than the boys. These statements are true of both our populations and Woodbury's

In *weight*, the girls also tend to be *relatively* more variable than the boys,—both in the case of our populations and Woodbury's However, for absolute variability in weight, the boys of Woodbury's groups tend to be slightly more variable than the girls, whereas, the girls of our populations tend to be slightly more variable than the boys.

Except then for the latter comparison, viz., absolute variability in weight, our results and Woodbury's are in agreement so far as trends of sex differences are concerned. For Garside's groups (4) the trend of sex differences in absolute variability is mixed. For relative variability in height, furthermore, there is no consistent trend of sex differences. However, in weight, the girls are, on the whole, consistently more variable than the boys This is also true of Baldwin's groups, previously cited (2), except for a marked difference at age 7 and slight differences at ages 8, 14, and 15, where the boys are the more variable in weight Pearson writes that girls "from 6 to 10 years of age are more variable than males in both weight and height" and that this "appears to be not a result of selection but of growth" (6) In a report of The Actuarial Society of America for 1929 (1, p 140), the data of tabulated results on relative variability in weight are described as confirming "the findings of Martin [1928, Germany] and Bardeen [1920, Carnegie Institute, Contributions to Embryology], quoted by Jackson [1929, American Journal of Physical Anthropology], who have shown that relatively greater weight for females is a general characteristic in nearly all races, ages, and statures"

Our results, as we have seen, are in accord with those just quoted for relative variability in weight And they are in agreement with Pearson's for both weight and height, although it should be emphasized that Pearson's statement about sex differences was for the period of 6 to 10 years of age If superior conditions of diet, sunshine, exercise, and rest, have any effect on sex differences, in relative weight and height variability, it is evident that such effect is not in the direction of a trend different from that ordinarily obtained If there is any effect, it probably is in the direction of accentuating the usual trend of sex differences However, the results of our population studies do not definitely reveal such accentuation when compared with the results of Woodbury's data

REFERENCES

1 ACTUARIAL SOCIETY OF AMERICA Medical Impairment Study New York, 1929

2 BALDWIN, B T The Physical Growth of Children *Univ Iowa Stud*, 1921, 1, No 1

3 BALDWIN, B T & WOOD, T. D Weight-height-age tables *Mother & Child*, 1923, 4.

4 GARSIDE, H V, ET AL Physical traits of young children *Amer J Dis Child*, 1929, **38**, 541-546

5. New York State Department of Health. Table of Average Heights According to Age Albany, New York, 1936

6 Pearson, K Data for the problem of evolution in man· III On the magnitude of certain coefficients of correlation in man, etc *Proc Royal Soc. London*, 1900, 66, 24-26

7 Peatman, J. G, & Higgons, R A Growth norms from birth to the age of five years A study of children reared with optimal pediatric and home care *Amer J Dis Child*, 1938, 55, 1233-1247

8 Woodbury, R M Statures and Weights of Children under Six Years of Age U S Dept Labor, Bur Pub. No 87, 1921

9 ——— Statures and weights of children under six years of age *Amer J Phys Anthrop*, 1922, 5, 5-16

Department of Psychology
College of the City of New York
New York City

Port Chester, New York

The Journal of Genetic Psychology, 1939, 54, 181-191

STUDIES IN MENTAL DEVELOPMENT: III PERFORMANCE OF TWELVE-MONTHS-OLD CHILDREN ON THE GESELL SCHEDULE, AND ITS PREDICTIVE VALUE FOR MENTAL STATUS AT TWO AND THREE YEARS*

Samuel S Fels Research Institute, Antioch College

Virginia Lafayette Nelson and T W Richards

As a continuation of the program initiated by investigating mental development of infants at six months, presented in two earlier papers (9, 10), the present discussion deals with a similar group at twelve months A total of 80 of the children in the Fels group were tested within five days of the twelfth month birthday, it was in connection with this group that data for the present report were obtained Of this group, about 60 per cent were present in the six months' group.

The percentage of Fels children passing each item, and similar percentages obtained with other children by four investigators are presented in Table 1 It will be seen that 17 items were passed by 90 per cent or more of the Fels group and that there were from two to five items in each of the remaining deciles That is, as in the case of the six months' schedule for this group, there was a fairly even distribution of items in terms of difficulty, except for a large lumping at the easier end of the scale. By comparing the percentage-passing obtained by the Fels group with the ranges given by Gesell (5), it is seen that 14 items were passed by more Fels than Gesell children, while eight items were found more difficult for Fels children Twenty-seven items were classified similarly in both studies. Comparing Fels percentages with Bayley (1) percentages, it is seen that six items differ by no more than five per cent, for three items the Bayley percentages are greater than the Fels figures, while for 11 items the Bayley percentage is less. The Linfert-Hierholzer study (7) compared with the Fels study shows seven items to be of about equal difficulty, four to be less difficult for the Washington children, and three more difficult. That is, the Fels group seemed

---

*Received in the Editorial Office on April 30, 1938

TABLE 1

Items of the Gesell Schedule Used at Twelve Months, with Numbers of Cases, Percentages Passing Each Item for the Fels Group, and Comparable Percentages for the Bayley, Fillmore, Linfert-Hierholzer and Gesell Groups

	Number cases	Fels Per cent pass	Bayley Per cent pass	Fill-more Per cent pass	Linfert-Hier-holzer Per cent pass	Gesell Per cent pass
Plays purposefully, cup, spoon, plate	81	100.0±0.0				85-100
Stands with help	80	98.8±0.8	80		100	85-100
Dangling ring pulls down	80	98.8±0.8	100			85-100
Adjust to words	81	98.8±0.8				85-100
Makes stepping movements	79	98.7±0.8	92		100	85-100
Inhibits on command	80	97.5±1.2	59			65- 84
Plays with blocks	81	97.5±1.2				65- 84
Secures cube	81	96.3±1.5				85-100
Ring bell, imitation	80	96.3±1.5				65- 84
Plays with mirror image	81	96.3±1.5	83			65- 84
Fine prehension	73	95.9±1.6	98			85-100
Imitates rattle spoon in cup	81	93.8±1.8	38		31	65- 84
Plays "peekaboo" or "patty"	80	93.8±1.8			94	85-100
Creeps	81	92.6±2.0	94		88	85-100
Paper: purposeful reaction	80	92.5±2.0				85-100
Walks with help	81	90.1±2.3	62	88		65- 84
Waves "bye-bye"	80	90.0±2.3				65- 84
Performance box: rod in hole	73	87.7±2.6				65- 84
Unwraps cube	80	87.5±2.5	80		59	50- 64
Puts cube in cup, play	81	85.2±2.7				20- 49
Puts cube in cup, imitation	78	81.6±2.8	90			85-100
Reaches for string	81	81.5±3.0				65- 84
Hits rubber doll	75	78.7±3.2	60			20- 49
Formboard, shown	73	76.7±3.4				65- 84
Inhibits hand to mouth	80	70.0±3.5	33			65- 84
Dangling ring uses string	81	69.1±3.5	100			65- 84
Scribble, imitation	81	69.1±3.5	87	64		65- 84
Accepts third cube	80	68.8±3.5		57		65- 84
Tower of two	80	60.0±3.7	20	37	38	50- 64
Says three words	80	60.0±3.7				85-100
Climbs	81	58.0±3.7			90	50- 64
Points to things at table	79	53.2±3.8			54	20- 49
Prefers one hand	81	48.2±3.8				65- 84
Vocalizes to music	79	48.1±3.8				65- 84
Says "bye-bye," "hello"	80	41.3±3.7			56*	50- 64
Tries to put on shoes	78	41.0±3.8				65- 84
Stands alone	81	40.7±3.7	37		84	20- 49
Bowel control	80	38.8±3.7			40	0- 19
Says four words	80	33.8±3.6				65- 84
Walks alone	81	32.1±3.5	31		30	20- 49
Formboard, circle	73	21.9±3.3				20- 49
Cup or plate	78	21.8±3.2				0- 19
Scribble spontaneously	81	21.0±3.1	17			20- 49
Says five words	80	18.8±3.0				0- 19
Uses spoon	80	16.3±2.8			44	20- 49
Tower of three	80	15.0±2.7	2			20- 49
Bladder control	80	8.8±2.2			18	0- 19
Asks for toilet	80	3.8±1.5				0- 19
Asks for things by name	79	3.8±1.5				0- 19

*"Bye-bye" only.

to be quite comparable with the group of Linfert and Hierholzer. With only five comparable items, it is seen that Fillmore's (4) children found three items more difficult and two of about the same difficulty. Thus, as in the case of six months' performance, the Fels group is seen to have no greater, and possibly less difficulty than the other groups.

A. THE TOTAL TEST AT TWELVE MONTHS

The items passed by each child were added to give a total score, as was done in the case of the six months' schedule. The distribution of the scores is given in Table 2.

TABLE 2

Score	Frequency
16.0 to 19.9	2
20.0 to 23.9	9
24.0 to 27.9	6
28.0 to 31.9	17
32.0 to 35.9	24
36.0 to 39.9	16
40.0 to 43.9	4
44.0 to 47.9	2
	80
Sigma	6.16
Mean	32.30

A reliability coefficient was obtained by using items in two series alternating in difficulty. The coefficient for the group was +.725±.03, which is raised to +.84 by the Spearman-Brown correction. This is lower than the +.88 reported by Linfert and Hierholzer and the +.93 reported by Fillmore for their groups at this age level, but higher than the +.76 reported by Bayley for her group. It is slightly lower than the (corrected) +.89 obtained by us at the six months' level.

B. SEX DIFFERENCES ON THE TWELVE MONTHS' GESELL SCHEDULE

Forty-one girls, with a mean total score of 33.8 were found to be slightly superior to 39 boys with a total score of 30.0, the critical ratio of the difference being 2.3 (no significant difference was apparent at six months). Considering the differences between the percentages of each sex passing each item, it was found that the following differences (critical ratios of two or more) favored the girls:

- Stands alone.
- Walks with help
- Walks alone
- Says four words.
- Tower of two
- Third cube.
- Performance box: rod in hole.

Boys excelled only on *scribbles spontaneously* and *dangling ring, uses string*

C. Relationship between the Performance on Single Items to Performance on the Total Test at Twelve Months

In Table 3 appear biserial coefficients between success or failure

TABLE 3

	Biserial r
Two-block tower	.88
Puts cube in cup, play	80
Puts cube in cup, imitate	76
Climbs	74
Hits rubber doll	.72
Says four words	72
Scribbles spontaneously	71
Performance box: rod in hole	70
Three-block tower	.66
Unwraps cube	65
Scribble imitation	64
Says five words	62
Tries to put on shoes	62
Says three words	61
Walks alone	59
Says "bye-bye" or "hello"	58
Takes third cube without dropping	56
Reaches for string	.52
Formboard shown	.51
Stands alone	51
Vocalizes to music	.40
Asks for things by name	.39
Inhibits hand to mouth	.36
Dangling objects, uses string	26
Prefers one hand	.26
Formboard: circle	.25
Cup or plate	.22
Bowel control	.21
Uses spoon	.16

on single items, and the total test score. (These coefficients are from 4 to 7 points higher than they would be if the effect of the particular item were eliminated in the correlation between it and the total.)

It will be seen that the coefficients range from .16 to .88. If we think of them in terms of r^2 and variance, it is apparent, thus, that single items could be regarded as incorporating the factors responsible for total test variation to the extent of from 2.6 per cent to 77.4 per cent. The median coefficient was .59, the mean .54, having r^2 values of .29 and .35.

Attempts to take out items by groups such, for example, as the locomotor-postural items or the language items failed to indicate what types of items better represent the total test than do other types. Thus, unlike the six months' Gesell Schedule which appeared to predominate in motor as opposed to alertness items, the twelve months' Gesell showed no obvious preponderance of any particular type.

D. THE RELATIONSHIP BETWEEN PERFORMANCE ON THE GESELL SCHEDULE AT TWELVE MONTHS AND PERFORMANCE ON TESTS AT LATER AGE LEVELS

Correlating her tests at 10-11-12 months, and 13-14-15 months with tests at 18-21-24 months and 27-30-36 months, Bayley obtained coefficients of .60 and .70 for periods most adjacent, and .45 and .54 for those most remote. Fillmore correlated her baby test total score at 11½ to 12½ months with an *IQ* presumably at two to three years and obtained a coefficient of .19. For single items in her series Fillmore obtained biserial coefficients with later *IQ* of as high as .47 (*piling some blocks*). Muhlenbein correlated 12 months' *LHQ*'s (Linfert-Hierholzer Quotients) and Stanford-Binet *IQ*'s at slightly less than five years. The coefficient of +.20±.12 she obtained was lower than that obtained by using the nine or six months' tests, though these also were low.

For intercorrelations between the total score on the Gesell Schedule at 12 months and later test performance the following indices were used: The total number of items passed on the Gesell Schedule at 18 months, raw score on the Merrill-Palmer at 24 months and at 30 months, and mental age on the Stanford-Binet (1916) at 36 months. Leeways in age were allowed as follows: At 18 months, seven days from the birthday, and one month on the Merrill-Palmer and Stanford-Binet tests.

Pearson coefficients obtained in these correlations are given in Table 4

TABLE 4

	r	r^2
Gesell 12 months with Gesell 18 months	.56±.07	.31
Gesell 12 months with Merrill-Palmer 24 months	.32±.08	.10
Gesell 12 months with Merrill-Palmer 30 months	.35±.09	.12
Gesell 12 months with Stanford-Binet 36 months (1916)	.33±.09	.11
Gesell 18 months with Merrill-Palmer 24 months	.47±.08	.22
Gesell 18 months with Merrill-Palmer 30 months	.50±.08	.25
Gesell 18 months with Stanford-Binet 36 months	.39±.10	.15
Merrill-Palmer 24 months—Merrill-Palmer 30 months	.84±.03	.71
Merrill-Palmer 24 months—Stanford-Binet 36 months	.80±.05	.64
Merrill-Palmer 30 months—Stanford-Binet 36 months	.77±.05	.59

It will be seen that the 12 months' Gesell correlated in no case very greatly with mental ability as measured in the tests at later levels It is seen in the r^2 column that 12 months' ability may account for but 31 per cent of variation at 18 months, and not over 10 per cent thereafter.

For purposes of comparison between six and twelve months' results correlations between six months' Gesell performance and later mental ability are given in Table 5.

TABLE 5

	r		r^2	
	6 Mo	12 Mo.	6 Mo	12 Mo.
18 months Gesell	.42	.56	.18	.31
24 months Merrill-Palmer	.37	.32	.14	.10
30 months Merrill-Palmer	.62	.35	.38	.12
36 months Stanford-Binet	.46	.33	.21	.11

It is seen that the six months' Gesell Schedule was correlated with later mental ability to a greater degree than was the 12 months' schedule, the earlier test may be considered to have sampled later mental ability more greatly than did the 12 months' tests.

By means of multiple correlation of the six months' and twelve months' Gesell Schedules with tests at 18 months (Gesell), 24 and 30 months (Merrill-Palmer), and 36 months (Stanford-Binet), the results were obtained that are given in Table 6.

In terms of r^2 it was found that the 12 months' Gesell increased

TABLE 6

	r Six months Gesell	r^2 Six and twelve months Gesell
12 months	72	
18 months	42	.56
24 months	.37	38
30 months	.62	64
36 months	46	46

the efficiency of prediction from the six months' Gesell alone for 18 months by 14 per cent, but for 24, 30, and 36 months by practically nothing. This is a remarkable finding and difficult of interpretation. One explanation may be that the 12 months' test is less reliable than the six months' tests ($r_{I\,II}$ of .84 as compared with 89), but this would seem insufficient to account for the difference above. A more plausible explanation is that the selection of items at six months affords a better all-around sample of ability tested at the later ages than does the 12 months' selection. Again, it is possible that the six months' series embraces a more even sampling of abilities required to pass the later tests, while the 12 months' test by selection of items is not so optimally weighted for these factors.

E. CORRELATIONS OF SINGLE GESELL ITEMS AT TWELVE MONTHS WITH THE MERRILL-PALMER TEST AT TWO YEARS AND THE STANFORD-BINET AT THREE

The distributions of Merrill-Palmer raw scores at two years and the Stanford-Binet mental ages at three years are given in Table 7.

TABLE 7

Merrill-Palmer Score	Frequency	Stanford-Binet *MA*	Frequency
5.0 to 8.9	3	34.0 to 36.9	13
9.0 to 12.9	1	37.0 to 39.9	7
13.0 to 16.9	17	40.0 to 42.9	6
17.0 to 20.9	14	43.0 to 45.9	1
21.0 to 24.9	13	46.0 to 48.9	12
25.0 to 28.9	10	49.0 to 51.9	2
29.0 to 32.9	2	52.0 to 54.9	3
33.0 to 36.9	0	55.0 to 57.9	1
37.0 to 40.9	1		
Total	61		45
Mean	20.05		42.49
Sigma	6.06		6.21

The Stanford-Binet distribution is thus seen to be bimodal. Coefficients of correlation (biserial) between 12 months' items and the Merrill-Palmer at two years are given in Table 8.

TABLE 8

	r
Says "bye-bye" or "hello"	.45
Vocalizes to music	45
Says three words	43
Points to things at table	33
Puts cube in cup, play	24
Says four words	19
Two-block tower	18
Hits rubber doll	17
Bowel control	17
Prefers one hand	15
Scribble imitation	14
Puts cube in cup, unit	13
Reaches for string	.13
Formboard shown	12
Performance box, rod in hole	06
Stands alone	06
Tries to put on shoes	05
Cup or plate	02
Five words	02
Uses spoon	—.01
Three-block tower	— 01
Inhibits hand to mouth	—.02
Climbs	— 02
Scribbles spontaneously	— 04
Formboard circle	— 05
Walks alone	— 05
Dangling objects uses string	— 05
Take third cube without dropping	— 09

It will be seen that the item correlations with mental ability at two years tend to be low, their median is 09 as compared with a median for the six months' Gesell of .25. Indeed, one item at the six months' level correlated .58 with the two-year Merrill-Palmer, while the highest in this series is 45 Multiple correlation between the first four of the items in the preceding table and the Merrill-Palmer at two yielded a coefficient of .615. This coefficient represents the combined effect of all items which correlate better with the Merrill-Palmer than does the total 12 months' test, so that it may be considered as an approximation of maximum correlation between the two tests Since R^2 of this coefficient is 3788, it may be inferred that less than 40 per cent of the variation at two years is sampled by the one-year test. This percentage is to be compared with about 50 per cent for six months' tests.

The data above suggest that the 12 months' Gesell Schedule, as a whole or considered item by item, affords no better prediction than does the six months' Schedule; indeed the predictive efficiency of the one-year test may be less. If the 12 months' schedule represents a good sampling of abilities at that age, it would seem that very little addition is made to the range of abilities between the periods from six to twelve months, at least of those abilities tested by the Merrill-Palmer Scale at two years. This does not mean, however, that the child at one year possesses abilities identical with those he possessed at six months. This is obviously not the case; indeed the correlation (.72) between the Gesell Schedules at six and twelve months would suggest that the repertoire of abilities sampled by them differs considerably. However, the two early tests do share much, and this portion seems to be tested to a fair degree at two years. The early Gesell, however, seems to possess more in common with the Merrill-Palmer at two years than does the one-year Gesell.

Biserial coefficients of correlation between 12 months' Gesell items and the Stanford-Binet at three years appear in Table 9.

These coefficients are seen to be smaller than the coefficients of items with the two-year Merrill-Palmer. The median coefficient is +.07. Three items correlate more highly with the Binet than does the total test. When the effect of these is combined they yield a multiple coefficient of .485 $R^2 = .2352$). The first two items—*says three words* and *scribble imitation* give a coefficient practically as high .480.

Comparing these results with similar results obtained for the six months' Gesell, it is seen again that the 12 months' test does not correlate as well with later mental status as does the earlier test. Results of the later test are evidence of no addition to the six months' repertoire of abilities—at least of those abilities tested later at two and three years. It is interesting to note, however, that of the tests which correlate best with later mental status there seems to be a preponderance of language items. Language items are not present to any extent in the six months' test. It would seem, therefore, that although language makes its initial appearance at 12 months and tends to correlate better with later mental status than do more motor tests, nevertheless, there seems to be no addition of fundamental

TABLE 9

	r
Says three words	.42
Scribble imitation	.37
Says "bye-bye" or "hello"	32
Points to things at table	.24
Bowel control	.24
Prefers one hand	21
Climbs	18
Hits rubber doll	14
Reaches for string	13
Dangling object uses string	12
Puts cube in cup, imit	10
Says four words	10
Vocalizes to music	09
Puts cube in cup, play	.08
Two-block tower	.07
Uses spoon	07
Inhibits hand to mouth	05
Cup or plate	.04
Three-block tower	04
Stands alone	04
Tries to put on shoes	04
Says five words	03
Performance box, rod in hole	02
Walks alone	00
Formboard circle	—.07
Scribbles spontaneously	—08
Takes third cube without dropping	—11
Formboard: shown	—16

ability with later mental status These beginnings of language are not indicative of a new fundamental ability but represent new indices, perhaps of such fundamental ability as is already present at six months

REFERENCES

1 BAYLEY, NANCY. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests. *Genet Psychol. Monog.*, 1933, **14**, 1-92

2 ———. The California first year mental scale. Berkeley, Calif.: Univ California Press, 1933 Syllabus Series No. 243. Pp. 24.

3 ———. The development of motor abilities during the first three years. *Monog Soc Res. Child Devel*, 1935, **1**, 1-26.

4. FILLMORE, EVA A. Iowa tests for young children. *Univ. Iowa Stud., Stud., Child Wel*, 1936, **11**, 1-58.

5. GESELL, A The Mental Growth of the Pre-School Child A psychological outline of normal development from fifth to the sixth year, including a system of developmental diagnosis New York Macmillan, 1925 Pp. x+447

6 JAFFA, A S The California preschool mental scale Univ Calif Press, 1934, Syllabus Series No. 251 Pp 66

7. LINTERT, HARRIETTE-ELISE, & HIERHOLZER, HELEN M A scale for measuring the mental development of infants during the first year of life *Stud Psychol & Psychiat*, 1925, 1, 1-33

8. MUHLENBEIN, JOSEPHINE N The validation of the Hierholzer Scale Unpublished Master's Thesis, Catholic Univ. of America, Washington, D C

9 NELSON, VIRGINIA L, & RICHARDS, T. W Studies in Mental Development. I Performance on Gesell items at six months and its predictive value for performance on mental tests at two years and three years *J Genet Psychol.* (In press.)

10. RICHARDS, T W., & NELSON, VIRGINIA L Studies in Mental Development II. Analysis of abilities tested at the age of six months by the Gesell Schedule. *J Genet Psychol* (In press)

11 STUTSMAN, RACHEL Mental measurement of preschool children with a guide for the administration of the Merrill-Palmer Scale of mental tests New York World Book, 1931 Pp. x+368

12 TERMAN, L M The Measurement of Intelligence An explanation of and complete guide for the use of the Stanford Revision and Extension of the Binet-Simon Intelligence Scale Boston Houghton Mifflin, 1916 Pp 362

Samuel S Fels Research Institute
Antioch College
Yellow Springs, Ohio

The Journal of Genetic Psychology, 1939, 54, 193-215

SOCIAL RELATIONS BETWEEN CHILDREN OF THE SAME AGE DURING THE FIRST TWO YEARS OF LIFE*

Department of Psychology, Fordham University

MARIA MAUDRY AND MARIA NEKULA[1]

A. THE PROBLEM

In her investigations on the first social reactions of the child, Charlotte Buhler (1) traced the development of the social relations observed in an experimental play situation between children of the same age through the first 18 months of life. There were indications in the material that even at this early age individual traits in the social behavior could be discovered. This led Charlotte Buhler to a tentative description of three types of social behavior, namely, the socially blind, the socially dependent, and the socially independent behavior (2).

The present study, too, was undertaken on the suggestion of Charlotte Buhler. Its original purpose was to investigate more in detail the individual differences; especially to find out whether the social attitude of a given child was consistent or changing from situation to situation.

In the course of the study, however, it soon became obvious that a more precise account of the development of the social relations during this period was a pre-requisite for any sound attempt to study individual differences. In her description of the "socially blind" child, e.g., Charlotte Buhler used, as a criterion, the fact that the child handles his playmate like play material, without any attempt at cooperation. This form of behavior varies decidedly with age. Therefore, it is necessary to be informed about the average amount of such behavior at the different age-levels before the statement can be ventured that a given child excels in it.

---

*Received in the Editorial Office on May 2, 1938.

[1] The junior writer was responsible for collecting the data used in this report, the senior writer supervised the research and prepared the manuscript for publication.

The first step taken in the treatment of the observational data was to establish the development of the social relations in the given situation. It is the preliminary report of this genetic part of our investigation that we present in the following pages.[2]

B. THE TECHNIQUE OF THE EXPERIMENTS

1. *The Experimental Situation*

Two children were placed together in a play-pen, which was situated in the middle of their nursery. After a period of about four minutes, during which no play material was present, three hollow cubes were distributed in the following way: each child was offered one cube, the third was placed on the floor between the children. After four minutes the cubes were taken away and a bell was given to one of the children. Again after four minutes the bell was taken away, a drum placed between the children, and a drumstick offered to each child. Four minutes later drum and sticks were removed and a ball was given to one of the children after they had been shown how to roll it from one to the other. The ball was left with the children for four minutes. Then the children were brought back to their cribs. The duration of the entire experiment was about 20 minutes.

The children were placed together by the experimenter, who left the room immediately afterwards, and took her place outside the glass walls of their nursery behind an observation screen. She entered the room only to present or change the play-material, or to rearrange the situation in cases of emergency. Cases of emergency were, e.g., when the children's fight became so intense that they were in danger, when a child fell or cried for some time, or when the children definitely turned away from each other.

2. *The Subjects*

As one of the purposes of the study was to investigate consistency in the social behavior of children, each of 24 children, henceforth referred to as the "Chief Subjects" was observed in four to twelve play situations with different children, hereafter called the "Partners."

---

[2] The discussion of the psychological literature on the topic is reserved for the final publication.

All the subjects were observed at the *Kinderubernahmestelle der Gemeinde Wien*, a social agency of the city of Vienna They were children of low social status who were temporarily in care of that institution.

Table 1 gives the age and number of chief subjects, the number of partners, and the observation time

TABLE 1

Age in months	Number of chief subjects	Number of partners	Observation time in minutes
6- 8	3	12	284
9-13	7	31	965
14-18	7	14	643
19-25	7	17	833
Total	24	68	2,725

The average time between the first and last observation of the same subject was 4 to 56 days with an average of 28 days With the exception of one child this time fell within the limits of the age groups of Table 1

The children were selected from the point of view of health as well as normal mental development Only children who readily adjusted to the social situation were taken as subjects As previous investigations had shown that a difference in age greater than two and a half months led invariably to the dominance of the older child and therefore limited the possible relationships, with a few exceptions the partner did not differ more than three months in age from the chief subject No children younger than six months were used as subjects, as Charlotte Buhler (1) had proved that active seeking of social contact does not occur before that age Furthermore only two children were placed together at a time since R Klein and E Wander (3) had shown that up to two years of age the child is not able to engage in cooperative play with more than one other child at a time.

3. *The Play Material*

The following play material was used

Three hollow cubes
A bell
A drum and two sticks
A ball

TABLE 2

First reaction	Inventory	Chief subject: Roland P , 1 , 9-3	Impulses	Partner. Erwin, B , 1; 10-23	Inventory	Fights and games
		Experimenter gives bell to Erwin				
Situation. material with partner First reaction turning to partner's material	N_1	sitting near rail, playing with beads, turns, looks at Erwin, grasps for bell	→			
			←	hides bell immediately behind back	N_2	Personal fight
	D_2	turns away, moves toward bar, moves beads back and forth	○			
			←	examines bell, watching Roland at the same time; rings bell	N_2	
	P_2	turns around to Erwin, babbles at him	→			
			○	deeply engaged in examining bell, pays no attention to Erwin	D_2	
		turns back to bar, moves beads to and forth				
		Experimenter gives drum and one stick to Roland, second stick to Erwin, takes away bell				
Situation material with both children First reaction turning to partner's material	A_1	grasps for Roland's stick, takes it away from him	→			
			←	takes stick back immediately hits drum, pulling drum to his knees	N_1	Blind fight
	C_2	hits drum	→			
	P_1	puts finger into his mouth, looks at Erwin	○ →		D_2	
			←	looks at Roland	P_1	
	C_2	hits drum, a few times, hits floor, hits drum again	→			

TABLE 2 (*continued*)

First reaction	Inventory	Chief subject Roland P, 1, 9-3	Impulses	Partner Erwin, B, 1, 10-23	Inventory	Fights and games
			O	puts stick away, turns drum upside down, examines drum, takes stick back	D_2	
			←	hits drum, looking at Roland and smiling at him	P_2	
	P_3	hits again, looking at Erwin stops hitting for a moment	→			
			←	looks at Roland, continues hitting	P_2	Personal game
	P_6	hits again, looking at Erwin	→			
			←	stops, looks at Roland	P_5	
	P_5	hits, looks at Erwin	→			
			←	hits again, looks at Roland, turns drum	P_2	
	P_1	stops hitting, sits still, watches Erwin	→			
			←	hits drum, smiles at Roland	P_4	
	C_2	looks at baby crying in the same room, puts finger in his mouth, hits drum with other hand, constantly looking at baby in crib	→			
			O	hits drum, stops hitting, looks at drumstick, takes drum to his knees, hits again	D_1	

For the children between six and nine months old, a rattle was used instead of the ball

On one side of the play-pen, three rows of wooden beads, movable on a horizontal wire, were fastened Although the children's attention was not called to the beads they at times became very interested in them especially when no other play material was available

The play material was selected from the point of view of its appropriateness for the age in question The period of four minutes without play material was used as a control situation

4 *The Method of Recording*

While the experimenter organized and supervised the situation to the extent necessary to establish and keep contact between the children, the observer took a detailed record of the behavior of both children. An example of part of a record as it was arranged for further analysis is given in Table 2 Right and left of the actual record the classification of the behavior of the children as discussed in the following sections is to be found.

As the record shows, the emphasis was put on the social behavior of the children. Less attention was paid to a detailed description of the behavior of the children when they were out of contact, or of a child's solitary play and manipulation of material. In the same way only that part of the experimenter's behavior was recorded which established or changed the situation.

Two limitations of the investigation must be mentioned.

(*a*) In the experimental situation the child finds himself confronted with a relatively unfamiliar child. The results of the study, therefore, do not give an account of all possible relations between children of the same age within the first two years of life, but of their relations in a special, though natural and important situation.

(*b*) The subjects, as shown above, are below average according to the social status of their parents; hence the possibility of retardation in social development must be taken into consideration

C. TREATMENT OF DATA AND RESULTS

1. *The First Reaction to the New Situation*

In the experimental situation the child found himself immediately confronted with play material and a playmate Extraneous factors were the other children and the furniture in the room. In this situation it was entirely left to the child where to turn, which opportunity to choose As the situation was rearranged at different times by the experimenter in the course of the experiment the observed child had to make his choice more than once (Table 2).

The first question that arises is Where does the child turn in a situation when he has a choice of play material, playmate, and other things around him?

A *new* situation was defined as every instance where the situation was rearranged by the interference of the experimenter, whether by giving new material, or by putting the children together This definition is based on the assumption, that by the interference of the experimenter the situation was so much altered that the reaction to it was a genuine choice. Only the reactions of the chief subjects were considered

By analysis, the first reactions of the child were grouped according to whether the child turned to the playmate, to the play material presented, or to the surroundings Social reactions were distinguished as positive and negative reactions. The material was classified into whether it was the partner's material, "free" material, i.e., material that was given to the child or not appropriated by the partner, or material on the partner, i e, dress, stockings, hair, etc. The distribution of the six possible choices according to age groups is given in Table 3.

The last column of the table shows that the interest in play material predominates, that in turning to the play material the child prefers the free play material which is more accessible, and only in second place turns to the partner's material This is the more important because the material was given more frequently to the partner than to the chief subject: out of 100 instances where material was given to only one child, it was given 57 times to the partner and 43 times to the chief subject Second in interest comes the playmate, who is welcomed by a positive address in about two out of three social reactions In about one-fifth of the reactions partner and test material are disregarded Turning to the material on the partner is negligible.

Apart from these general results important developmental trends are shown in the changes of the distributions of these first reactions with the different age-groups. Turning to the surroundings as well

TABLE 3

FIRST REACTION IN NEW SITUATION

Child's first reaction is directed	Numbers Age in months					Percentages Age in months				
	6-8	9-13	14-18	19-25	All ages	6-8	9-13	14-18	19-25	All ages
To free material	39	193	114	149	495	19.5	30.9	25.5	23.4	25.9
To partner's material	56	155	101	103	415	28.0	24.8	22.6	16.3	21.7
To material on partner	6	7	2	0	15	3.0	1.1	0.5	0.0	0.8
Positively to partner	33	77	100	210	420	16.5	12.4	22.5	33.0	22.0
Negatively to partner	5	66	39	62	172	2.5	10.5	8.7	9.7	9.0
To surroundings	61	129	91	112	393	30.5	20.3	20.4	17.6	20.6
Total	200	627	997	636	1910	100	100	100	100	100

as turning to the partner's material show a steady decline to about half their original frequency, which is most pronounced between the first and second age-groups[3]. The former indicates an increasing interest in the play situation, the latter may be interpreted as a growing respect for the partner's property, which is accompanied by a definite increase in positive reactions to the partner from 13 months on. Turning to the free material and positive reaction to the partner show development which is significantly related: the former increases between 9 and 18 months and shows a steady decline afterwards. The latter between 9 and 13 months and shows a steady increase afterwards. Furthermore the ratio between the total of turning to the test material and the total of social reactions is 2.7, 2.4, 1.6, and 0.9, for the successive age-groups. Both facts may be interpreted as a shift of interest from the material to the partner. The negative reactions to the partner co-vary somewhat with the interest in free material, showing an equally pronounced increase between 9 and 13 months followed by a smaller and less regular decrease afterwards. Considering the ratio between positive and negative reactions to the partner it can be said, that the social attitude is a negative one between 9 and 13 months, a positive one before and afterwards.

Further explanation of this development, especially of the change of social attitude, was possible by taking into consideration the placement of material. To investigate its influence the child's reactions were tabulated separately for the following situations present in the experiment: test material with chief subject, test material with partner, test material with both children, no test material.

Table 4 shows the distribution of the first reactions of the chief subjects in relation to the placement of the material.[4] Since contact with the playmate and play with the test material are the two special opportunities which the experimental situation offers to the child, the table may be interpreted in the following way.

---

[3] The differences between these proportions have been found to be statistically significant using as criterion a ratio of 2 as critical, viz., a difference was considered significant if $\frac{dp}{\sigma_{dp}} > 2$

[4] The blank spaces indicate the lack of opportunity for the stated reaction in the given situation, e.g., in the situation without test material there is no possibility for turning to the partner's or to free material.

TABLE 4

FIRST REACTION IN RELATION TO PLACEMENT OF MATERIAL, IN PERCENTAGES

Chief subject's reaction is directed	Situation without test material Age in months				Test material with partner Age in months				Test material with chief subject Age in months				Test material with both children Age in months			
	6-8	9-13	14-18	19-25	6-8	9-13	14-18	19-25	6-8	9-13	14-18	19-25	6-8	9-13	14-18	19-25
To free material									48.8	52.8	43.3	42.2	39.4	47.1	37.7	35.4
To partner's material					58.6	69.9	63.1	39.4					24.2	24.5	18.5	15.0
To material on partner	15.8	5.3	3.3	—	—	1.0	—	—	—	—	—	—	—	—	—	—
Positively to partner	42.1	44.7	50.0	56.0	17.3	3.4	10.7	28.0	2.4	4.7	18.9	28.9	6.1	9.6	21.6	28.0
Negatively to partner	—	—	5.0	2.7	1.3	15.9	11.7	9.3	7.3	18.9	12.2	11.4	3.0	10.5	8.0	12.2
To surroundings	42.1	50.0	41.7	41.3	22.6	8.8	14.5	23.3	41.5	23.6	25.6	17.5	27.3	8.3	14.2	9.4
Total	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100

Between six and eight months a response to the experimental situation is quite unspecific Although a positive social attitude prevails in the behavior of the child, it reveals no special interest for the partner. This is shown by the relatively small number of social contacts, especially in the situation where the chief subject or both children have material; by the fact that in a situation without material the child turns as often to the surroundings as he does to the partner, by the fact that only in this age group the child turns to material on the partner instead of to the partner himself Neither is there a special interest in the test material: this is shown by the fact that in situations where the chief subject has the material, he turns nearly as often to the surroundings as to the test material; that furthermore most of the positive reactions occur in the situation where all the material is with the other child This is the situation where the negative attitude is most expressed later on.

Between 9 and 13 months a specific reaction to the experimental situation develops and the test material is the opportunity to which the child responds first. The partner does not yet receive special attention. This is shown by the fact that turning to the test material is at its maximum in all situations, that turning to the surroundings shows a noticeable increase and predominates over the social reactions in the situation without material In connection with this interest in play material, the attitude toward the partner changes. In all situations where material is present the negative reactions to the partner are at their maximum, the positive at their minimum A negative attitude prevails This negative attitude, however, is not due to a genuine hostility, but to the predominant interest in material, to the acquisition of which the partner is often a hindrance It is noticeable that there is no negative reaction to the partner in the situation without material, and nearly as much positive as negative responses to the partner when both children are provided with material.

From 19 months on the child responds definitely to the social opportunity of the situation The interest in the playmate partly replaces the interest in material. This is shown by the fact that the positive reactions to the partner form nearly a third of the responses, that they by far exceed the negative ones and are more frequent than turning to the surroundings; that this positive attitude for the first time is present in all situations and thus has become independent of the placement of material, that a decided drop is noticeable in turning to the partner's material in the situation where all material is with him, indicating a growing respect for the partner's property. Up to that age material was even more interesting when the partner had it than when it was in the child's possession.[5]

The period between 14 and 18 months can be described as a transitional phase. Under "favorable" circumstances, i.e., when both children have material, or when no material is present, the child is able to display the positive attitude and interest in the partner characteristic for the next age group. In "unfavorable" circumstances as, e.g., when only the partner has material he falls back into the negative attitude characteristic for the previous age group.

2. *Inventory of Behavior*

In the preceding section the first reaction to the experimental situation was analyzed, i.e., the attitude a child developed upon meeting a partner and play things in a new situation. The next problem of interest is the analysis of the behavior of the child throughout this situation, i.e., an inventory of his social responses.

To subject the records to quantitative treatment, it was necessary to divide them into units of behavior Units or "impulses" were defined as items of behavior that show the same social tendency, in the records they are marked by arrows. As the division into units of behavior was made from the social point of view they were not necessarily single movements, but might be complex reactions consisting of different movements accompanied by expression (Table 2). The impulse might be answered by an impulse of the other child or the other child might disregard the impulse. In the latter case it was marked zero Here again only the social attitude of the child was considered. A zero was recorded whenever a new impulse of the partner was not answered regardless of whether the child was engaged in different preoccupations (Table 2) By analyzing the impulses and the zero reactions we could devise the following systematic inventory of response occurring in the experimental situation

---

[5] Ratio between turning to partner's material in situation where only partner has material and turning to own material in situation where only chief subject has material.

a Disregard of interference (*D*).

(1) *Lack of interest in Partner*

*D*1), taking no notice of partner, looking around, turning away, looking at nurse, experimenter, or other child

*D*2), taking no notice of aggression of partner, turning away or to other material without sign of resistance or fear

(2) *Absorbed in material*

*D*3), deeply engaged in play with own body or material so that partner's address or his aggression remains unnoticed.

b Chance contact and interference without social intention (*C*).

*C*1), physical contact or touching partner by chance, taking material lying near the partner without paying attention to the latter

*C*2), manipulating, looking at, grasping for, moving towards the same material simultaneously with partner without paying attention to the latter

*C*3), manipulating parts of partner's body or dress with or without material, looking only at the part attracted No signs of paying attention to the partner

*C*4), noticing partner's activity and taking it up, without paying further attention to him

*C*5), taking material given to him, thrown away or put aside by partner, looking at material in the partner's hands without paying attention to the latter

c. Negative social behavior (*N*)

(1) *Fight for material.*

*N*1), taking material from partner or trying to get it, after defeat watching partner closely until he abandons material and then seizing it

*N*2), drawing own material nearer, covering it with hand for protection, moving it away from partner, holding it tightly, withdrawing it.

*N*3), watching partner or material closely after fight with signs of tension

(2) *Personal aggression*

*N*4), hitting, pushing, biting partner, pulling his hair, addressing him with sounds of anger

*N*5), disapproving of partner by making faces, crying, screaming

d Positive social behavior (*P*)

(1) *Looking and grasping*

*P*1), looking at partner or his movements

*P*2), smiling, babbling, at partner

*P*3), moving towards partner, touching him with or without material, grasping for him, caressing him while paying attention to his reactions

(2) *Cooperation in play*

*P*4), attempt to start cooperation by showing things, performing, imitating, giving or taking in friendly manner with attention to reaction of partner

*P*5), cooperation, both children engaged in the same activity while paying attention to each other, mutual adjustment of activity to each other

In devising this system the social significance of the different items was used as a criterion

Group I, disregard of interference, comprises all items of behavior that lead to an interruption of the contact established This might occur as a result of mere lack of interest in the partner or because the interest in material was stronger

Group II, interference without social intention, comprises all items of behavior where the contact of the children contains no sign of social interest and can be explained entirely by chance or the fact that both children were independently interested in the same thing.

Groups III and IV comprise the items of behavior where the child is paying attention to his partner.

Group III, negative social behavior, consists of the hostile reactions. A differentiation is made, whether they occur in the course of attempting to obtain or defend play material or whether it is a case of personal hostility

Group IV finally comprises all forms of friendly contact between the children Division is made between more primitive manifestations of interest in looking, smiling, grasping at each other and more complex reactions, where mutual adjustments are a pre-requisite.

In the sample record (Table 2) the impulses and reactions are marked in the middle column. The groups to which they belong are marked right or left of each impulse in such a way that the letters correspond to the main groups, and the numbers to the subgroups of the systematic inventory

Table 5 gives the distribution of the different impulses for the different age groups

TABLE 5
BEHAVIOR DURING THE EXPERIMENT

Age in months	Interference without social intention	Negative social reactions	Positive social reactions	Total number of impulses	Disregard of partner's addresses	Total number of reactions
6- 8	336	393	363	1,092	847	1,939
9-13	723	2,001	1,382	4,106	2,355	6,461
14-18	347	1,349	1,347	3,043	1,611	4,654
19-25	435	1,460	1,810	3,705	1,676	5,381

TABLE 6
DISREGARD OF PARTNER'S ADDRESSES, IN PERCENTAGES OF TOTAL REACTIONS

	Age in months			
Disregard for	6-8	9-13	14-18	19-25
(*a*) Lack of interest	29.3	21.1	19.3	15.7
(*b*) Absorbed in material	14.2	15.4	15.3	15.5
Total	43.5	36.5	39.6	31.2

Table 6 shows the frequency of disregard of the partner's address. Up to eight months nearly half of the addresses of the partner are disregarded; from nine months on, this number decreases with the most pronounced difference between the first two age groups. If the instances of disregard are divided into (*a*) disregard for lack of interest in the partner and (*b*) disregard for absorption in the play material, it becomes obvious that the decrease is entirely due to changes in the former group, indicating the growth of social responsiveness

Figure 1 takes up the children's interference with each other, and shows the distribution of impersonal, negative, and positive contacts at the different age groups

Between six and eight months positive and negative social reactions and impersonal interference occur with about the same frequency Between 9 and 13 months there is a decrease in the chance contacts and a corresponding increase in the negative social reactions, the latter are the outstanding form of contact at this age level and at

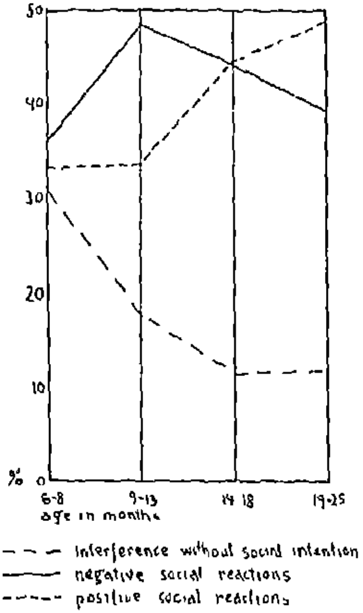


FIGURE 1
INTERFERENCE DURING EXPERIMENT

the same time attain their maximum frequency for the entire age range. Throughout the ages between 14 and 25 months there is an increase in positive social reactions, due mainly to a further decrease of chance contacts (14 to 18 mos.) and of negative social

TABLE 7
SOCIAL BEHAVIOR, IN PERCENTAGES

Impulses	Age in months 6-8	 9-13	 14-18	 19-25
Interference without social intention	30.9	17.6	11.4	11.7
Fight for material	33.1	45.3	35.9	29.8
Personal aggression	2.8	3.4	8.4	9.6
Total of negative social reactions	35.9	48.7	44.3	39.4
Smiling, looking, grasping	32.5	28.3	32.4	25.1
Cooperation	0.7	5.4	11.9	23.8
Total of positive social reactions	33.2	33.7	44.5	48.9
	100	100	100	100

reactions (19 to 25 mos) Between 19 and 25 months positive relations have become the most important form of contact. Table 7 gives a further analysis of the negative and positive social reactions, according to the groups (*a*) and (*b*) in the systematic inventory.

It is noticeable that throughout the observed period fight for material is by far more frequent than personal aggression, that only fight for material shows the above discussed rise and decline of the negative reactions, whereas personal aggression increases steadily

Of the positive social reactions looking and grasping for the partner is present from the beginning and furnishes about 30 per cent of the responses through all the age groups The changes are entirely due to a development of more complex forms of contact In other words, the progress consists in increasing ability to cooperate in play, which develops in addition to the existing primitive forms of contact, as grasping, looking, and smiling

These findings may be interpreted in the following way From the age of nine months on, up to which the attitude towards the partner is unspecific, showing an equal amount of friendly, aggressive, and chance contacts, a definite negative attitude develops It consists chiefly in fights for material and is most obvious between 9 to 13 months After 18 months, it changes into a positive attitude, due to the development of cooperation in play Throughout the development there is a steady increase in response to the social opportunities of the situation which manifests itself in the decrease of chance contacts, and of disregard of partner, as well as in the increase of personal aggression and cooperation

3. *Fight and Play*

An outstanding feature in the development of social relations is the parallel decrease of disregard for the partner's address and of the chance contacts, it leads to the conclusion that with increasing age the behavior becomes more integrated, the contact closer and the situation more organized.

As this conclusion cannot be proved by analysis of the single impulse, it seemed advisable to investigate the development of more complex behavior, taking as units coherent situations In the experimental situations given "games" and "fights" were the appropriate units. A "game" was defined as the total of at least two impulses through which the same activity or activity with the same object is carried out undisturbed by both children; the single impulses constituting the game consist either in positive social reaction or in activity with the same object; however the occurrence of an isolated negative reaction or disregard of the partner's actions in the course of the situation is not considered as ending the game.

Similarly, "fight" was defined as the total of at least two impulses of negative social behavior. Again an isolated occurrence of positive address, impersonal interference or disregard was not considered as ending the situation.

Using these criteria it was possible to single out games and fights occurring in the course of the observation period.

The games and fights were analyzed according to their frequency and duration, the nature of the play-activity, and the amount of attention paid to the partner.

Table 8 shows the frequency of fights and games as measured by their occurrence per 10 minutes of observation and their length as measured by the number of impulses.

TABLE 8

Frequency and Length of Fights and Games

		Fights				Games			
		Age in months							
		6-8	9-13	14-18	19-25	6-8	9-13	14-18	19-25
	Number of impulses per 10 minutes	3.2	3.8	3.3	3.1	0.7	0.6	0.8	1.9
		0/0	0/0	0/0	0/0	0/0	0/0	0/0	0/0
Length of situation	2 impulses	50.0	36.5	41.3	41.6	73.7	58.6	56.3	41.5
	3-4 impulses	34.4	34.8	33.2	32.7	26.3	30.0	32.8	33.7
	5 or more impulses	15.6	28.7	25.5	25.7	—	11.4	10.9	24.8
		100	100	100	100	100	100	100	100

The following results are of interest. Fights are more numerous throughout the investigated age groups, their frequency is rather constant with a slight maximum between 9 and 13 months. Games show a constant frequency up to 18 months and thereafter a remarkable increase.

Fights are longer than games during the first and at the beginning of the second year, by the end of the second year games and fights show equal length. The maximum of length coincides in both cases with the maximum of frequency

From the point of view of frequency as well as length, between 14 and 18 months the fights show the characteristics of the next older age groups and the games those of the next younger age group, in other words, the significant decrease in fights occurs between 14 and 18 months, the significant increase in games after 18 months of age.

In investigating the amount of attention the child paid to his partner during the fights and games, a differentiation was made between indirect and direct contact. The contact is indirect if both children are playing near each other or with the same object, or trying to get hold of the same object without paying attention to each other This is called "blind play" or "blind fight" In the case of direct contact the friendly or hostile reactions are directed first of all at the partner. Both children try to influence directly the partner's behavior, and modify their own reactions in order to have a mutual play situation or to get the desired object This is called "personal play" or "personal fight."

The third possibility is that one of the children plays or fights blindly and the other is aware of the social relation involved. This is considered a transitional stage.

Table 9 shows in percentages the distribution of these three types of play and fight at the different ages.

In both situations at the beginning of the period investigated, fights and games are almost entirely blind, whereas at the end at least one

TABLE 9

TYPES OF FIGHT AND PLAY ACCORDING TO AMOUNT OF ATTENTION PAID TO THE PARTNER

Situation	Age in months 6-8	9-13	14-18	19-25
Blind fight	90	69 4	44 5	42 9
Transitional stage	10	27 5	40 9	31 7
Personal fight	—	8.1	14 6	25 4
Total	100	100	100	100
Blind play	100	54 3	25	15 7
Transitional stage	—	11 4	14	7 2
Personal play	—	34 3	61	77 3
Total	100	100	100	100

of the children fights and both of the children play with full attention to the partner in the majority of cases. Fights throughout the development remain blind to a higher extent, showing once more that interest in material and not personal hostility is their direct course. The interrelation of active interest in the play material and social attitude is further shown by the fact, that the games become predominantly personal at the same time that the first drop of fight for material occurs (Table 7).

Finally it is of interest to analyze the use which is made of the material in the play situation. By differentiating the activities that occurred within the play situation according to their social significance the following three groups were found: I. Handling the same object, II. Social use of material; III. Primary social reactions.

Group I: Handling the same object, contains all random manipulation such as sucking, shaking, hitting, rubbing, tearing things, etc., as well as all manipulations where proper use is made of the play material, such as beating the drum with the drum sticks, ringing the bell, putting the hollow cubes in and out of each other, etc.

In both instances the child does not give up the material though he may adjust his manipulation to that of the partner.

Group II. Social use of material implies activities where the material is not of primary interest but becomes a means to establish and maintain contact. The appeal the social situation has for the child is strong enough to make him relinquish his playthings for the sake of contact with the other child. Activities belonging to this group are: showing, giving things, rolling ball to the other child, alternate exploration of material or parts of body and dress, etc.

Group III. Reactions which are purely social, without material being an essential means of contact, such as directing the other child, helping, consoling, begging, showing off.

TABLE 10

TYPES OF PLAY ACTIVITY, IN PERCENTAGES

Activity	Age in months 6-8	9-13	14-18	19-25
Handling same object	100	88.9	57.6	56.6
Social use of material	—	11.1	38.0	37.6
Purely social reactions	—	—	4.4	5.8
	100	100	100	100

Table 10 shows the distribution of the different play activities at the different age levels

It indicates that throughout the investigated period the play consists chiefly in handling the same material Parting from one's material for the sake of social contact begins after nine months of age and from 14 to 25 months stays practically the same after a definite increase

A comparison of the results of Tables 9 and 10 shows, that the characteristic change of the games as far as attention to the partner and social use of material is concerned takes place between 14 and 18 months. Between 19 and 25 months the progress consists chiefly in that the previously established way of playing together becomes more frequent.

D Summary and Conclusion

Ninety-two children from 6 to 25 months were observed in standardized play situations with a partner not more than three months older or younger within the given age limits

To trace the development of social reactions within this period, the records were analyzed from three points of view· (*a*) first reaction in the new situation, (*b*) inventory of reactions to partner during the situation, (*c*) fights and games in the course of the experiment

The following paragraphs try to integrate the results by describing for each age level the child's response to the two main opportunities which the experimental situation offers to him, i e , social contact with the playmate and play with the test material. An attempt is made to derive from this the different rôle of the partner in the course of development

Six to eight months The child does not yet respond to the opportunities of the situation One out of three times he turns immediately to the surroundings (Table 3) If he enters the situation the reactions to the material and to the partner are equally unspecific. The newly introduced play material gets little attention unless it is given to the partner (Table 4). Yet this does not call forth a negative attitude to the partner. For in the same situation most of the friendly approaches occur (Table 4) and throughout the situation there are as many positive contacts as negative ones (Figure 1). On the other hand nearly half of the addresses of the partner are disregarded mostly for lack of interest in him (Table 6) Interference with the partner is as often without social intention as it is a positive or negative social reaction (Figure 1). The friendly contacts are limited to looking, smiling at, and grasping for the partner (Table 7), i.e., except for the smile the partner is treated about the same as play material. The games are few and short (Table 8) consisting in unspecific manipulation of the same object (Table 10). They are blind, i.e., the partner does not receive special attention (Table 9) Fights are equally impersonal, consisting in a blind attempt to get hold of the material (Table 9).

Nine to 13 months: The play material is the opportunity, to which the child responds first. As the partner often becomes an obstacle to its acquisition, a negative attitude dominates whenever material is present (Table 4) Fighting is at its maximum (Table 8), it has become the most frequent behavior throughout the situation, and has partly subplanted distraction and impersonal interference (Figure 1, Table 7). It furthermore has become personal on the part of at least one child to a considerable degree (Table 9). Yet this is due to the increasing interest in material and not to the development of a genuine hostility; the child prefers free material (Table 3) The negative attitude is not present in the situation without material and is least pronounced when the desire for material is satisfied as, e g , when both children have material (Table 4). The general tendency of a more personal relation is noticeable also in the friendly contact, the first instances of cooperation and social use of material occur (Tables 7, 10), and the playmate receives special attention for the first time (Table 9).

Nineteen to 25 months: The child realizes the opportunity for social contact and makes use of it by integrating his social interest with his interest in play material. Although at first sight the material is still more attractive than the partner (Table 3), a friendly attitude towards the latter prevails independently of the placement of material (Table 4) Positive social reactions predominate throughout the situation (Figure 1), in which cooperation has become as important as the more primitive forms of looking, smiling, and grasping (Table 7), games show a considerable increase in frequency and length (Table 8) The play is personal, i e., both children modify their behavior in adjustment to the partner's activity (Table 9), the material is no longer a hindrance but a means for establishing social relations as social use of it has become an important feature of the game (Table 10).

Fourteen to 18 months In this transitional period the shift of interest from the material to the partner is indicated by the predominance of a friendly attitude to the partner in "favorable" situations, i e., when the child's desire for material is satisfied (Table 4). It is also shown in a pronounced decrease of fighting for material, increase of cooperation (Table 7), and the fact, that the games acquire the structure characteristic of the next age group, as far as attention paid to the partner and social use of material are concerned (Tables 9, 10)

Throughout the age range the rôle of the partner, as characterized by the outstanding feature of the play situation, is successively that of play material (6-8 months), obstacle to play material (9-13 months), and playmate (19-25 months)

REFERENCES

1. BUHLER, CH Soziologische und psychologische Studien uber das erste Lebensjahr *Quel. u Stud. Jugenkun.*, 1927, **5**
2. ——— The social behavior of children In *Handbook of Child Psychology* (ed C Murchison) Worcester Clark Univ Press, 1933.
3. KLEIN, R, & WANDER, E Gruppenbildung im zweiten Lebensjahr. *Ztsch f Psychol*, 1933, **128**, 257-280

Department of Experimental Psychology
Fordham University Graduate School
New York City

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1939, 54, 217-221

COMMENTS ON RECENT STUDIES OF *VTE*

Department of Psychology, University of Virginia

WAYNE DENNIS AND R. W. RUSSELL

In many of the reports on the nature of the learning process in the white rat, which have been published since 1900, attention has been called to a type of behavior occurring at a point in a discrimination box or in a maze at which the animal must make a choice of pathways. This behavior has been characterized by Hoge and Stocking (2, p. 48) as a tendency on the part of the animal to "... look to right and left several times before making a choice." Lashley (3, p. 327) has described the phenomenon as follows: " . . . the beginning of discrimination was marked by long hesitation with a swaying back and forth between passages." Furthermore, it has been noticed that an animal might enter a passage but turn about and reverse its course without proceeding to the end of the alley.

These phenomena, while commonly observed, have received no generally accepted terminology but have been labelled variously as head movements, head "vibrations," starting errors, partial errors, partial entrances, partial eliminations, "lookings back and forth," and "runnings back and forth."

Recently Muenzinger (4, 6) has proposed the additional term "vicarious trial and error" (abbreviated to *VTE*) for " . . . a facing into one alley before the other one, whether right or wrong, was entered . . . ." In further describing this behavior (4, p. 77) he has stated that it may be accomplished in various ways.

> The most common way is for the rat to stop at a mid-point between the alleys and turn his head first towards one and then towards the other alley. But he may also approach the entrance to the alley and orient his whole body towards it and then turn and approach the other alley in a similar way.

Muenzinger and Gentry (5, p. 204) had previously referred to this "choosing" behavior as including a tendency to a " jerk towards the correct alley and a jerk away from the wrong alley ." before entering one of the alleys These statements appear somewhat ambiguous as regards the exact pattern of *VTE* behavior and are not specific with reference to whether slight movements of the head into the wrong pathway are to be considered as errors or as *VTE* The distinction between an approach, orientation, or "jerking" toward an alley and partial entrance into the alley seems to be merely one of degree.

Muenzinger notes that behavior which he calls *VTE* has been reported by numerous experimenters, but he states that it was Tolman in 1926 (8, p 367) who first suggested that these "lookings back and forth" or *VTE* movements *might be related to learning efficiency* To be sure, it would appear that the idea might have been original with Tolman, for in his book of 1932 he suggests that someone study the relation of "lookings back and forth" to learning (9, p. 216). This would imply that Tolman was not acquainted at the time with the work of this character already performed.

Muenzinger, himself, has failed in his general survey of the relation of *VTE* to learning efficiency (4) to credit earlier workers with observations of a like nature Particularly since his general paper is apparently the first in a series on *VTE*, we are disturbed by the statement that the relation of head movements to learning is a new and unworked problem, and we feel that in justice to previous investigators it is important that attention should be called to the existence of data on this problem in the earlier literature The subsequent discussion is not intended to detract from the significance of Tolman's or Muenzinger's observations, but is presented merely to place their contributions in chronological relationships to some of the earlier results

In an article by Lashley (3) on visual discrimination in the white rat which appeared in 1912 the author reported that

> " in the early stages of learning, a correct choice more frequently followed where the animal hesitated than when it made a rapid choice. Later in the experiment choice by negation seemed to become the more important method of eliminating error. The rats chose a passage quickly, and if it were wrong, turned back when about eight inches from the stimulus. Choice by affirmation never seemed to be very effective, even with the best trained rats. They would frequently make a correct choice, then hesitate, turn back, and compare the forms" (3, p 327)

Here, then, is a very concise statement of relationships between what is called *VTE* and other behavior.

A monograph by Peterson in 1917 (7) contains many data on the degree of entrance into maze alleys at different periods in the course of learning. In this study, the least degree of error which was systematically recorded was called a "start," by which it was meant that only the head or the fore part of the body entered an alley, the rat then pivoting and entering the alternative pathway. This degree of entrance may be more extensive than that which Muenzinger has referred to as *VTE*, but the difference between starts and *VTE* would seem to be merely one of extent of movement. In addition to starts, Peterson also recorded "half-entrances," which were defined as all entrances greater in length than starts and less than complete entrances.

Peterson found that starts were most numerous just before the period at which all errors cease, a finding which Muenzinger has reported for *VTE* behavior in certain discrimination problems (4). Peterson's monograph shows graphically in Figures 4 and 5 that while complete entrances are at their greatest frequency on the first trial and decrease rapidly, starts and half-entrances begin with a low frequency and increase for a period as the complete errors decline in number.

Peterson discussed head movements as well as starts, calling the former "head vibrations." He wrote as follows (7, p. 28):

> In numerous cases when the habit of avoiding the *cul de sac* was nearly complete, the peculiar rapid vibration of the head noted above took place. The pause was, however, but for an instant. This response seems to indicate that the impulse to go forward at the critical place is still partly checked or impeded by one to enter, not quite eliminated.

Peterson also referred to the relation of head movements to learning efficiency in his conclusions as follows:

> The elimination of entrances to blind alleys does not come about mainly by a decrease in the *number* of entrances, but principally, especially in the case of the longer *culs de sac*, by a gradual decrease in the *degree*, or the distance, of entrance. Just before entrance is eliminated completely, there frequently occurs a peculiar and very rapid vibration of the rat's head between the direction of the true path and that of the tempting blind alleys.

This also is directly analogous to the findings recently reported by Muenzinger

These statements led the senior author (1) of the present paper to pay special attention to head movements in a study of maze learning in the white rat which appeared in 1930 The mazes were single-unit *T*-mazes, but the three mazes which were used required the rat to run different distances between the point of choice and the goal. All head movements at the point of choice were recorded, as were partial entrances

In this study head movements were found to be related to learning as follows head movements occurred more frequently in the second half of the trials required to meet the criterion than in the first half (This is in accord with the findings of both Peterson and Muenzinger.) Furthermore, the trials on which head movements occurred were correct in a much larger percentage of the instances than were the trials which preceded and which followed the trials in which these movements were noted This fact shows a very intimate connection between head movements and accuracy in individual runs up to the point at which the maze is run without error

In Dennis' study a further result was obtained which has not as yet been subjected to test by others. The longer mazes, in which a wrong choice meant the traversal of a very long blind alley, elicited many more head movements than did the short maze In the short maze the *cul de sac* was six feet long, whereas in the other two mazes the *cul de sac* was 12 feet and 24 feet respectively Choosing the latter blind meant the traversal of 48 feet of excess pathway, and the head movements which occurred as the rat faced this choice were much more frequent than when the blind was shorter. This may be related to Muenzinger's finding that rats in a discrimination problem exhibit more instances of *VTE* when shocked than they do when no shocks are given, as in this case also the choice involves more consequences to the animal

A perusal of Muenzinger's paper will show that nearly all of his findings have been in some measure anticipated by others while, as noted earlier, he implies that the relation of head movements to learning is an unexplored problem There are, of course, new features in the experimental conditions which Muenzinger employed. Furthermore, part of the data here reviewed have come from maze studies whereas Muenzinger's data are from records of discrimination learning. However, we noted previously the general observations of Lashley on the topic of the relation of head movements to the learning of visual discriminations.

In the matter of the interpretation of head movements Muenzinger is more bold than other writers, with the possible exception of Tolman. As is shown by the name which he gives to the phenomenon, Muenzinger argues that these movements of the animal are a form of vicarious trial and error. We are certainly in no position to deny this claim, nor, we think, is it possible for anyone to establish it. What is the definition of "trial and error behavior" and in what sense are the head movements vicarious? What would be the distinction between vicarious trial and error and *incipient* trial and error, *inhibited* trial and error, or *vestigial* trial and error? What findings are necessary in order to prove that head movements are *VTE* in some other sense than by definition? It appears to us that these questions must be answered before Muenzinger's hypothesis can be subjected to a test.

REFERENCES

1. DENNIS, W. A study of learning in the white rat. *J. Genet. Psychol.*, 1930, **37**, 294-308.
2. HOGE, M. A., & STOCKING, R. J. A note on the relative value of punishment and reward as motives. *J. Anim. Behav.*, 1912, **2**, 43-50.
3. LASHLEY, K. S. Visual discrimination of size and form in the albino rat. *J. Anim. Behav.*, 1912, **2**, 310-331.
4. MUENZINGER, K. F. Vicarious trial and error at a point of choice. I. A general survey of its relation to learning efficiency. *J. Genet. Psychol.*, 1938, **53**, 75-86.
5. MUENZINGER, K. F., & GENTRY, EVELYN. Tone discrimination in white rats. *J. Comp. Psychol.*, 1931, **12**, 195-205.
6. MUENZINGER, K. F., & FLETCHER, F. M. Motivation in learning. VI. Escape from electric shock compared with hunger-food tension in the visual discrimination habit. *J. Comp. Psychol.*, 1936, **22**, 79-91.
7. PETERSON, J. The effect of length of blind alleys on maze learning: an experiment on twenty-four white rats. *Beh. Monog.*, 1917, **3**, No. 4. Pp. 53.
8. TOLMAN, E. C. A behavioristic theory of ideas. *Psychol. Rev.*, 1926, **33**, 352-369.
9. ———. Purposive Behavior in Animals and Men. New York: Century, 1932. Pp. 463.

Department of Psychology
University of Virginia
University, Virginia

The Journal of Genetic Psychology, 1939, 54, 223-227.

"PERSISTENCE" TEST SCORES OF STUDENTS COMPARED WITH RESPECT TO THE NATIVITY OF THE MALE PARENT*

William Woods College

DAVID G. RYANS

The psychology of individual differences is of relatively recent origin and its records are, as yet, conspicuously superficial and incomplete It has been only within the last four decades that the psychologist has shown more than "passing" interest in the very evident, and significant, fact that performances of individuals in similar situations vary widely one from another But today the "normal adult mind" is no longer considered to be the *sole* subject matter of psychological investigation, thoroughgoing knowledge of the deviations from the average is also deemed necessary to an understanding of behavior.

With the recognition of trait differences in psychology came a natural interest in the differentials or variables with which the degree of possession of a given trait appeared to be related Thorndike early pointed out certain factors to which various physical and behavioral differences might be traced. *Race*, *age*, *sex*, and ultimately, of course, *nature* and *nurture* were suggested as variables of causal importance in the study of the existence and measurement of individual differences Many traits were observed to vary, more or less significantly, with regard to averages, from one race or nationality to another. For example, differences in "intelligence," as tested, seemed apparent between the *white* and *yellow* race groups, on the one hand, and the *black* on the other Speed of reaction, psychopathic condition, ascendance-submission, introversion, honesty, and countless other traits, likewise, were reported as showing variations with respect to the racial factor. While many of the results possessed low reliability from a statistical standpoint, suggestive trends were indicated

Recently (1) an effort was made to analyze situations which seemed to demand persistent behavior, or continued energy release,

---

*Received in the Editorial Office on March 4, 1938

on the part of a responding individual. As an extension of this study a "persistence test" was constructed (2). The test, demanding (*a*) consecutive effort at rational learning, (*b*) self-ratings on a "functional" schedule, and (*c*) physical endurance, appeared to possess sufficient validity and reliability to permit its use in the study of human behavior

In light of the apparent measurability of "persistence," the question of differences in this capacity between ethnic groups becomes a pertinent one On the average, do differences in perseverance between members of different races or nationalities exist? If so, to what extent; which groups seem to be more and which less persistent as revealed by methods of behavior sampling? It is believed that the results to be reported here, while obviously preliminary, have some bearing on such questions

In gathering the essential data for the investigation, the "Persistence Test" was administered to 442 high school and junior college students Their educational backgrounds were approximately similar, in that all of their homes were in the same community or its environs

No attempt will be made to state accurately differences due to *race* as the term *race* is technically employed Most authorities agree that there are perhaps three primary racial stems which can be differentiated upon the basis of physical characteristics (And in our present state of knowledge it is not feasible to refer to classifications based upon behavior traits Indeed it seems probable that if behavior differences do exist, function and structure vary *dependently*) The most acceptable classification of primary races includes the *yellow, white,* and *black*. In the absence of Persistence Test scores on members of the yellow and black races, our results are limited to the white groups. And here, again, it has not been possible to follow a strict anthropological classification and divide our subjects according to the daughter races (Nordic, Mediterranean, or Alpine) of the white race to which they belong This has not been possible, in the first place, because there is extensive overlapping of basic features and the physical characteristics are by no means distinct In the second place, the shifting of European boundaries has made the type of information which was available almost impossible of satisfactory analysis Therefore, the comparisons of Persistence Test scores which are attempted are with respect to the birth-place of the father of the subject, the groups are admittedly linguistic and political rather than anthropological. If an individual's male parent were born in Italy, for example, the subject was classed as an Italian. Cases of international marriage, of course, invalidate our results, but, insofar as marriage outside one's own national group is the exception rather than the rule with the parents of the samples studied (subjects were predominantly first generation Americans, of foreign-born parentage) this factor is not considered to be of great effect.

Thirteen different nationalities have been classified in Table 1 with respect to scores made on the Persistence Test. In all, 21 nationalities were represented by the sample studied, but groups of which there were only one, two, or three members were omitted in these tabular comparisons. The national groups which were made up of from 25 to 122 members are shown apart from those of smaller membership in the table.

TABLE 1

COMPARISONS OF THE PERSISTENCE TEST SCORES OF 442 HIGH SCHOOL AND JUNIOR COLLEGE STUDENTS CLASSIFIED WITH RESPECT TO NATIVITY OF THE MALE PARENT

Group	N	Mean	SD	% cases above Q_3 of total distribution*	% cases below Q_1 of total distribution**
(A) National groups represented by from 25 to 122 cases					
Finnish	70	152.9	20.7	31.9	24.7
American	47	152.5	14.8	33.6	14.7
Italian	57	151.0	18.2	26.6	21.0
English***	25	150.7	16.2	32.0	20.0
Swedish	43	146.3	19.2	13.8	29.9
Slovenien	122	145.7	17.0	17.2	28.7
Croatian	26	142.5	15.7	11.7	31.2
(B) National groups represented by from 5 to 12 cases					
Norwegian	9	162.0	18.8	44.0	11.0
Serbian	11	158.9	16.4	45.0	18.0
Austrian	6	156.6	12.6	34.0	00.0
Jewish	5	147.5	6.3	00.0	00.0
German	12	147.1	16.7	24.9	33.2
Irish	9	144.2	12.5	11.0	22.0

* Q_3 of total distribution = 161.0

** Q_1 of total distribution = 137.0

*** English includes Scotch

In general, it appears that the Finnish, American, Norwegian, Serbian, and Austrian nationalities rank higher in the possession of persistence, while the Croatian, Slovenien, and Irish are substantially lower.

For the most part the national differences here observed are statistically insignificant The nearest approaches to significant differences among the larger samples, *i.e.*, those represented by 25 or more cases, appear to exist between the means of the Finnish and Croatian, and American and Croatian groups. In both instances the critical ratio is 2.6, which indicates that the differences are just short of being reliable. The difference between the averages of the Finns and Croatians is 10.4 points and the *SD* $_{(Diff)}$, 3.97 points Between the American and Croatian samples the difference between the Means amounts to 10.0, whereas the standard deviation of the difference is 3.79. When the smaller samples (section *B* of table) are considered along with the larger ones, the greatest apparent differences exist between the Norwegian and Croatian and the Serbian and Croatian groups. In the case of the former, the difference between the Means is large, 19.5, but the *SD* $_{(Diff)}$ is correspondingly large, being 6.99 The averages of the Serbian and Croatian groups, likewise, differ by 16.4 points and the standard deviation of the difference is 5.82. In both of these instances the reliability of the difference falls short of the figure assumed to represent a "true difference," the obtained *D/SD* $_{(Diff)}$ being 2.8.

More extensive testing is desirable before conclusions may be validly drawn

SUMMARY AND CONCLUSION

Four hundred and forty-two high school and junior college students were classified upon the basis of the birthplace of the male parent and were studied with respect to their scores on a "Persistence Test." Among the groups studied there seemed to be a tendency for Finns, Americans, Norwegians, Serbs, and Austrians to excel in persistent behavior, and for Croatians, Irish, and Slovenicns, to possess the capacity in lesser amounts as compared with a number of other groups The results are to be regarded as tentative

REFERENCES

1. RYANS, D. G. An experimental attempt to analyze persistent behavior. I Measuring traits presumed to involve "persistence." *J. Gen. Psychol.*, 1938, **19**, 333-353

2. ——— An experimental attempt to analyze persistent behavior: II A persistence test. *J. Gen. Psychol.*, 1938, **19**, 355-371

William Woods College
Fulton, Missouri

The Journal of Genetic Psychology, 1939, 54, 229-234

A TENTATIVE STATEMENT OF THE RELATION OF PERSISTENCE TEST SCORES TO CERTAIN PERSONALITY TRAITS AS MEASURED BY THE BERNREUTER INVENTORY*

William Woods College

DAVID G RYANS

In earlier discussions (2), (3), results were reported which, it is believed, indicated the possibility of providing quantitative determinations of persistent behavior, or the organism's *capacity for continued energy release* Measures of performances presumed to involve "persistence" on the part of the individual were obtained, tables of inter-correlations set up, and a multiple factor analysis attempted A general, or common, factor, which was contributed to by several of the original test situations and which was decidedly influenced by ratings on perseverance, seemed evident. Further experimentation led to the development of a "Persistence" Test Comprised of three sections measuring (*a*) continued effort in a learning situation, (*b*) self-estimation, on a "functional" scale, and (*c*) endurance of the cumulative effects of muscular fatigue, the test appeared to be sufficiently valid and reliable to justify its use as an indicator of persistent behavior

It may be of interest to take note of some of the relationships into which this so-called capacity of "persistence," as measured by the Persistence Test, enters with personality determinations indicated by the six scoring keys commonly applied to the Bernreuter Personality Inventory.

The sample with respect to which comparisons reported here are made is a small one, consisting of 40 junior college sophomores who had made up one of the experimental groups in the original analysis of persistence The sexes were equally represented Ages ranged from 18 to 25 years There was no intentional selection. All subjects had attended the same secondary school The range of scores on the Revised Army Alpha Examination extended from 80 to 203, on the Otis Self-Administering Test, Higher Examination, from

---

*Received in the Editorial Office on April 16, 1938

28 to 69. The Mean Alpha score was 148.0, with a standard deviation of 21.15, while the Mean for the Otis Test was 47.65 and the standard deviation of the distribution 8 16. On the Persistence Test the Average score was 154.9; the standard deviation, 22 5.

The personality indices to be considered here were obtained through the administration of the Bernreuter Inventory and its subsequent scoring with the six keys presumably measuring (*a*) neurotic tendency, (*b*) self-sufficiency, (*c*) introversion-extroversion, (*d*) dominance-submission, (*e*) confidence in self, and (*f*) sociability. The scales are described by Bernreuter as follows:

> B1-N A measure of neurotic tendency Persons scoring high on this scale tend to be emotionally unstable Those scoring above the 98 percentile would probably benefit from psychiatric advice Those scoring low tend to be very well balanced emotionally
>
> B2-S. A measure of self-sufficiency Persons scoring high on this scale prefer to be alone, rarely ask for sympathy or encouragement, and tend to ignore the advice of others. Those scoring low dislike solitude and often seek advice and encouragement.
>
> B3-I. A measure of introversion-extroversion. Persons scoring high on this scale tend to be introverted, that is, they are imaginative and tend to live within themselves. Scores above the 98 percentile bear the same significance as do similar scores on the B1-N scale Those scoring low are extroverted; that is, they rarely worry, seldom suffer emotional upsets, and rarely substitute day dreaming for action
>
> B4-D. A measure of dominance-submission. Persons scoring high on this scale tend to dominate others in face-to-face situations Those scoring low tend to be submissive.
>
> F1-C A measure of confidence in oneself. Persons scoring high on this scale tend to be hamperingly self-conscious and to have feelings of inferiority; those scoring above the 98 percentile would probably benefit from psychiatric or medical advice Those scoring low tend to be wholesomely self-confident and to be very well adjusted to their environment.
>
> F2-S A measure of sociability. Persons scoring high on this scale tend to be non-social, solitary, or independent. Those scoring low tend to be sociable and gregarious. (1)

In Table 1 product-moment coefficients of correlation obtaining for Persistence Test results and Bernreuter Inventory scores are

TABLE 1

PRODUCT-MOMENT CORRELATION COEFFICIENTS INDICATING THE EXTENT TO WHICH THE BERNREUTER PERSONALITY INVENTORY SCORES OF FORTY JUNIOR COLLEGE STUDENTS VARIED WITH RESPECT TO PERSISTENCE TEST PERFORMANCES AND INTELLIGENCE INDICES

Scale		Persistence test score		Intelligence index	
		r	*P.E.*	*r*	*P.E.*
B1-N	Neurotic tendency	—0.19	0.11	0.20	0.10
B2-S	Self-sufficiency	—0.14	0.11	0.02	0.11
B3-I	Introversion	—0.27	0.10	0.20	0.10
B4-D	Dominance	—0.01	0.11	—0.10	0.11
F1-C	Lack of confidence	—0.14	0.10	0.05	0.11
F2-S	*Non*-sociability	0.07	0.11	0.08	0.11

given. In the right-hand column, the coefficients between typical "intelligence" indices (based upon the equation of Army Alpha and Otis Test results by application of standard deviation techniques) and Bernreuter scores are shown for purposes of comparison. None of the correlation coefficients appear, on the surface, to be significant. The most probable relationships, however, concern *lack* of neurotic tendency and persistence, tendency toward *extroversion* and persistence, and *self-confidence* and persistence.

Further analysis of the data (and some verification of the major trends noted in the preceding paragraph) is provided by Table 2. This table embraces a comparison of the highest and lowest fourths of a Persistence Test score distribution from the standpoints of the various Bernreuter scores. In only two instances do differences appear which approach statistical significance. There seems to be a slight tendency for the more persistent individuals to be more stable emotionally, while the less persistent members of the sample show Bernreuter scores indicating increased tendency toward neuroticism. A similar inclination is observed in comparisons which are made with respect to introversion-extroversion, the less persistent tend to be "introverted" whereas the more persistent individuals lean toward the extroversion end of the scale. A *Diff./S.D. (Diff.)* of 1.6 would seem to indicate the bare possibility that the proneness of subjects who are more persistent to be, at the same time, more confident of themselves is a reliable trend.

TABLE 2

COMPARISONS OF THE MOST PERSISTENT FOURTH AND THE LEAST PERSISTENT FOURTH OF A SET OF PERSISTENCE TEST SCORES FROM THE STANDPOINTS OF TRAITS MEASURED BY THE BERNREUTER PERSONALITY INVENTORY

	Average neurotic scores	S.D	$\frac{D}{SD_{(diff)}}$	% scores above P.R. of 50	% scores below P.R. of 50
	1. Neurotic tendency				
Highest quarter of persistence distribution	54.0	16.6		40.0	60.0
Lowest quarter of persistence distribution	72.5	21.6	2.1	90.0	10.0
	2. Self-sufficiency				
	Average sufficiency score				
Highest quarter of persistence distribution	49.5	28.4		50.0	50.0
Lowest quarter of persistence distribution	42.5	30.4	0.5	40.0	60.0
	3. Introversion-extroversion				
	Average introversion score				
Highest quarter of persistence distribution	41.0	11.2		20.0	80.0
Lowest quarter of persistence distribution	67.5	24.2	3.1	80.0	20.0
	4. Dominance submission				
	Average dominance score				
Highest quarter of persistence distribution	44.0	21.0		30.0	70.0
Lowest quarter of persistence distribution	39.0	25.8	0.5	40.0	60.0

TABLE 2 (Continued)

	Average neurotic scores	S.D	$\frac{D}{SD_{(Diff)}}$	% scores above P.R. of 50	% scores below P.R. of 50
	5. Lack of confidence				
	Average lack of confidence score				
Highest quarter of persistence distribution	61.0	15.2		70.0	30.0
Lowest quarter of persistence distribution	75.0	23.1	1.6	90.0	10.0
	6. *Non* sociability				
	Average lack of sociability score				
Highest quarter of persistence distribution	56.0	28.6		70.0	30.0
Lowest quarter of persistence distribution	58.5	32.7	0.2	60.0	40.0

SUMMARY

When the Persistence Test scores of 40 junior college sophomores were studied with respect to personality traits measured by the Bernreuter Inventory it was observed that (*a*) negative product-moment correlation coefficients, ranging from 0.01 to 0.27, were obtained in all cases except that when "sociability" was involved, (*b*) these measures of concomitant change were, for the most part, substantiated by comparisons of extreme quarters of the Persistence Test distribution, and (*c*) the trends indicated the possibility of a relationship existing between "persistence," on the one hand, and emotional stability, extroversion, and confidence in self, on the other. The results are to be considered preliminary.

REFERENCES

1. BERNREUTER, R. G. Manual for the Personality Inventory. Stanford University, Calif.: Stanford Univ. Press, 1935.

2. RYANS, D G. An experimental attempt to analyze persistent behavior I Measuring traits presumed to involve "Persistence" *J Gen Psychol*, 1938, 19, 333-353

3 ———, An experimental attempt to analyze persistent behavior II A persistence test *J Gen Psychol*, 1938, 19, 355-371

William Woods College
Fulton, Missouri

APPARATUS

The Journal of Genetic Psychology, 1939, 54, 235-236

GROOVED BLOCKS FOR USE IN REGISTERING CHOICE

Cambridge, Massachusetts

MARIE L H FORBES

The drawing (Figure 1) shows a rectangular block with grooves for solid rubber type, the block being one of a collection of blocks to be used ready-inked in printing, with a fixed frame, on tape.

Since the type-face assembled on the base of the block describes but does not reproduce the caption on the top of the block, the latter is not limited to color and configuration. Samples of materials flavored or scented, rough or smooth in texture, or common objects such as hooks, nails, or screw-eyes may be attached to the blocks. The experimenter displays an item to be matched or calls for it by name, the subject responding by placing the required block in the printing frame instead of in the hand of the experimenter

A series of blocks which can be manipulated by the normal child, the mentally defective child, and the ape, with type-faces which print on tape a record intelligible to the experimenter only, can be imagined as of use in a comparative study

The writer gratefully acknowledges the examination of the materials by her former instructor, Dr Walter F. Dearborn, Director of the Psycho-Educational Clinic, Harvard University

---

*Received in the Editorial Office on March 17, 1938

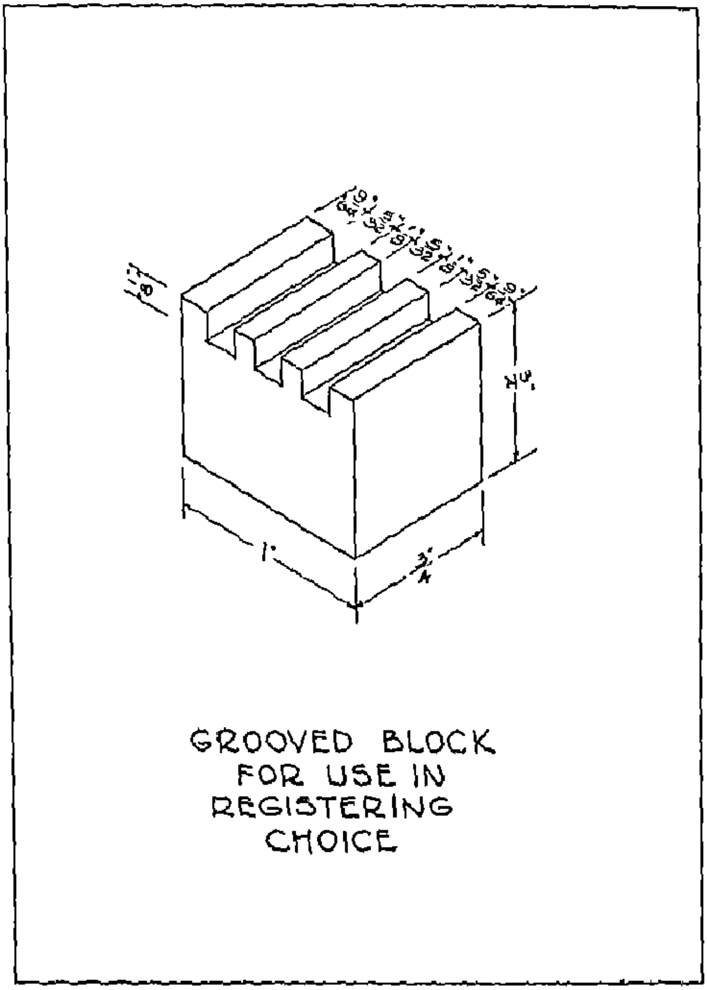


FIGURE 1

8½ Ash Street Place
Cambridge, Massachusetts

BOOKS

The Journal of Genetic Psychology, 1939, 54, 237-262

If a book is to be adequately discussed, every opinion concerning it must be given publicity, and recurring opinions must be registered as such. That is, a controversial book must be reviewed many times, and not just once or twice We speak here of reviews at the professional level, not of childish blurbs or vindictive pecks.

The *Journal of Genetic Psychology*, the *Journal of General Psychology*, and the *Journal of Social Psychology*, will buy reviews in the open market at not less than $2 per printed page and not more than $3 per printed page.

Conditions Only those books that are listed below in this section are eligible for such reviews. All general elementary textbooks are eliminated Those books that are included, even though some of them might be poor, are probably the ones that make the most difference to psychology or deal with problems or techniques that are fundamental to psychology at the present time The list begins with January, 1936, and the books are listed approximately in the order of receipt. New 1938 books will be added as they come in. At the end of 1938, all of the 1936 books will be removed from the list

A reviewer must possess the Ph D degree or its equal in training and experience. The Editor thinks of graduate students with affectionate regard, but he cannot accept reviews of this type from them.

A review must be written strictly at the professional level It must not occupy less than two printed pages, nor more than twenty. It must avoid trivialities, such as chapter divisions, spelling of words, typographical errors, or any other matters that the reader is not looking for in that book A review must not be a soap-box or pulpit from which the reviewer exhibits his own intellectual hobbies or private feelings. It is the book that is being reviewed, and the book must occupy the stage No competent review will use such make-believe weapons as *"but this is not psychology,"* or *"but this is biology."* The classifications of old-fashioned college catalogues are not of vital importance in these columns. A competent book deals with issues, or with techniques for the investigation of issues A competent review identifies those issues, determines their importance, and evaluates the success or failure of the book in the accomplishment of its purpose. For a reviewer to point out that the purpose of a book is not his purpose, is in fact a statement by means of which the reviewer substitutes himself for the book. That type of vulgarity has no place on the stage of great books

Procedure. If among the books below there is one about which you have thought, and concerning which you have formulated some ideas, you are invited to write a review of that book No matter if the book has been reviewed several times or a dozen times. It is important that your colleagues know your comprehension of the book and its significance. In this way professional opinion will prevail quickly, and uncritical theories, unimportant or badly conceived investigations, misinterpretations, insufficient evidence, or uninspired work of any kind will stand revealed.

(Authors of reviews will receive a check immediately the size of which will be within the limits named above. Nominations for inclusion in this list will be received gladly by the Editor)

(1937)

Allport, G W. Personality: A Psychological Interpretation New York. Holt, 1937. Pp. 588.

Beck, S. J Introduction to the Rorschach Method. *Res Monog, Amer Orthopsychiat Asso.*, 1937, No 1.

Newman, H. H., Freeman, F N, & Holzinger, K J. Twins A Study of Heredity and Environment Chicago Univ Chicago Press, 1937. Pp 369.

Murphy, G, Murphy, L B, & Newcomb, T M Experimental Social Psychology. New York: Harper, 1937 Pp 1121.

Southall, J P C Introduction to Physiological Optics New York and London. Oxford Univ Press, 1937 Pp. 426.

Terman, L. M., & Merrill, M. Measuring Intelligence Boston: Houghton Mifflin, 1937. Pp 461.

Thomas, W I. Primitive Behavior: An Introduction to the Social Sciences New York: McGraw-Hill, 1937 Pp. 847.

Stagner, R. Psychology of Personality. New York McGraw-Hill, 1937 Pp. 465.

Murphy, L. B Social Behavior and Child Personality. New York Columbia Univ. Press, 1937. Pp 344

Anastasi, A. Differential Psychology New York. Macmillan, 1937. Pp. 615.

BURROW, T. The Biology of Human Conflict. New York: Macmillan, 1937. Pp. 435.

GESELL, A., & ILG, F. L. Feeding Behavior of Infants. Philadelphia: Lippincott, 1937. Pp. 210.

(1938)

THURSTONE, L. L. Primary Mental Abilities. *Psychomet. Monog.*, 1938, No. 1. Pp. 121.

SPEARMAN, C. Psychology Down the Ages. (2 vols.) New York: Macmillan, 1938. Pp. 454 & 355.

GESELL, A., & THOMPSON, H. The Psychology of Early Growth. New York: Macmillan, 1938. Pp. 290.

STRONG, E. K. Psychological Aspects of Business. New York: McGraw-Hill, 1938. Pp. 629.

BOAS, F. The Mind of Primitive Man. (Rev. ed.) New York: Macmillan, 1938. Pp. 285.

SHERMAN, M. Mental Conflicts and Personality. New York: Longmans, 1938. Pp. 319.

STEVENS, S. S., & DAVIS, H. Hearing: Its Psychology and Physiology. New York: Wiley, 1938. Pp. 489.

GUTHRIE, E. R. The Psychology of Human Conflict. New York: Harper, 1938. Pp. 408.

MURPHY, G., & LIKERT, R. Public Opinion and the Individual. New York: Harper, 1938. Pp. 316.

KATZ, D., & SCHANCK, R. L. Social Psychology. New York: Wiley, 1938. Pp. 700.

ELLIS, W. D. A Source Book of Gestalt Psychology. New York: Harcourt, Brace, 1938. Pp. 403.

WOODWORTH, R. S. Experimental Psychology. New York: Holt, 1938. Pp. 889.

SKINNER, B. F. The Behavior of Organisms: An Experimental Analysis. New York: Appleton-Century, 1938. Pp. 457.

SEASHORE, C. E. Psychology of Music. New York: McGraw-Hill, 1938. Pp. 408.

TERMAN, L. M. Psychological Factors in Marital Happiness. New York: McGraw-Hill, 1938. Pp. 474.

MURRAY, H. A., *et al.* Explorations in Personality. New York: Oxford Univ. Press, 1938. Pp. 761.

LANDIS, C., & PAGE, J. D. Modern Society and Mental Disease. New York: Farrar & Rinehart, 1938. Pp. 190.

MUNN, N. L. Psychological Development. New York: Houghton-Mifflin, 1938. Pp. 582.

RASHEVSKY, N. Mathematical Biophysics. Chicago: Univ. Chicago Press, 1938. Pp. 340.

CRITICAL REVIEWS OF RECENT BOOKS

(*Stevens, S Smith, & Davis, Hallowell. Hearing Its Psychology and Physiology New York, Wiley, 1938 Pp. 489.*)

REVIEWED BY MAX F MEYER

Since the scope of a scientific book must have its limits, and one should probably hesitate to criticize the authors for not having done what they regard as lying outside, the reader may try to conclude what he might or might not expect, from the dedication, which in this case is to two Harvard professors: (*a*) one because of his "constant concern for auditory theory," (*b*) the other because of his "experiments on the electrophysiology of the nervous system"

The book begins as a textbook on the college level, introducing the reader into that physical terminology of acoustics which has grown up during the last two decades. This task is usefully done. Probably there are more than a few older psychologists, physiologists, otologists, who have with some bewilderment heard of such things as phons, electrical-mechanical analogies, vacuum tube oscillators, sound filters. They may here clarify and modernize their thinking

In the succeeding chapters the authors have diligently collected, *e g ,* all the recent determinations of the intensity threshold in various aspects, including the amplitude of the eardrum, and the calculation of the amplitude of the interior of the cochlea Then all the modern experiments on the dependence of pitch perception, not only on frequency, but also on other factors, especially on intensity, are collected from the variety of scattered sources.

Occasionally something phantastic is given the reader to look at, for instance, Figure 32 on page 91, which refers to "the basilar membrane" and states positively that such a disturbance as pictured "will there be produced," while the authors down on the same page and the following pages show and admit that they have no clear conception at all of what happens in that membrane This is far from enlightening the unsuspecting reader Hypothesis and experimental fact should each be distinctly marked as such, and the particular "ifs" of the hypothesis should not be forgotten in statements made

The authors report the case of the missing fundamental as if it were something new The fundamental is not there and yet it is distinctly heard. Is that a mystery? They (and the telephone engineers who absent the fundamental by filtering it out) entirely fail to see that this is an old phenomenon, for a hundred years utilized by the organ builders when the money for a very long pipe is not available, but a very low tone is wanted, and they do not observe that this and the difference tone are exactly the same thing, and that it points to the very theory of the motion of the basilar membrane which the authors do not regard as even worth mentioning. We return to this later. But let the reviewer add here that a "fundamental" is really nothing fundamental, but merely the largest common factor in all the frequencies, and as such a difference of them. Then the mystery has a chance to disappear. But the book is too uncritical for that.

Some experimenting, and some more theorizing, has been done of old, and still more of them recently, about the threshold of successiveness. The authors introduce us to the more recent discussions of this type. Loudness and masking are well presented by them, without any effort, however, to solve on their own part some of the innumerable puzzles still left, which are as amazing as the planetary motions were before Kepler. Much work done in recent years on the loudness of (brief) tones as depending on duration, important in telephone practice, is reported by the authors. Volume, density, brightness, are discussed as tonal attributes. Auditory localization as dependent on time, phase, intensity, is represented by numerous quotations; also the stereophonic effect. Combination tones are described, but not recognized as being not always what by the authors they are supposed to be. (Rücker and Edser, not mentioned, showed 50 years ago that the middle ear apparatus is irrelevant for the hearing of difference tones in most situations, and many others have brought the same proof since that time,—with no effect on the authors.) Fatigue and persistence are described. The authors justly criticize the "flicker" method of investigating persistence. In optics that is feasible; in acoustics not, because here the flicker, by Fourier's law (or theorem), only introduces new frequencies.

The authors, although they are still largely thinking in terms of resonance in the cochlea, have to admit against their will that persistence (which in resonators would be unavoidable, of course) has not been found in the cochlea or the auditory nerve as great as 0.14 sec. This time, mentioned, seems to be that of decay, but, they add, it must be a central mechanism. They fail to add: "and this proves that resonance is not essential for the functioning of the cochlea." However, they admit on page 203 that "Ohm's Law" which presupposes resonators in the cochlea, is untrue. Nevertheless, most chapters of the book rely on it.

The authors draw an interesting conclusion (p. 240) about the vibrato. The reviewer (be he permitted to confess this weakness!) regards the vibrato as a distinctly unesthetic, *i.e.*, unpleasant, aspect of music, and some psychologists of the University of Iowa on the contrary have hailed it as one of the greatest contributing factors of musical esthetics. The book says: "It covers up errors—if a vibrato singer goes off key, the audience is unaware of it."

In the middle of the book the anatomy of the ear is well described. The illustrative figures are very good. More should have been said about the mechanical properties of the cochlea. The microscope shows that the basilar membrane is fastened ("anchored" and not inserted, the reviewer asserts) to the sides in such a manner that a static tension is out of the question, and a relatively very great (hydraulic) displacement must be necessary before any considerable tension results; these (to the reviewer patent) facts, the latter especially, are left unmentioned, nor are the consequences presented. Also Ter Kuile's important suggestion that the (microphonic) distortion of the hair cells is probably due to a rather definite kind of mechanical interaction between the hairs and the tectorial membrane is left unmentioned, his name does not appear in the book. All along, the book belittles every contribution which dates farther back than to be found in the acoustic periodical literature of the last half dozen years, the reader should keep that in mind before he is willing to accept this (as offered by the publisher) as an "authoritative" book. The resonance theory is dogmatically introduced into the anatomy on page 273, and no equal favor of plausibility through anatomy is given any other mechanical view. Yet page 278, an after-thought seemingly, adds an "as if" and moreover refers to Appendix II, about which a word later here. Conclusions "from models" are blackballed; yet on page 285 pictures are reproduced to show in detail how the authors think "waves travel" in the cochlea in spite of the fact that these pictures were a drawn model and not observed experimentally on the animal cochlea itself. It all seems to depend on whether the "model" pictures are made by a personal friend. For that same reason, probably, one cochlea model by Békésy is given an unlimited passport on another page.

After plowing through two-thirds of the book we come to green fields. The neurophysiology is taken up, and here we have indeed an excellent treatment of the subject matter Piezoelectric effects are distinguished sharply from action potentials of a cell Distortion of a living cell causes the former and leaves the cell virtually unweakened and free from any refractory period. This is very plausibly shown to be the function of any and all of the hair cells of the cochlea. It is called the microphonic function of the hair cell. On the other hand, the nerve cell is exhausted by its action potential wave like a storage battery and must immediately be recharged The result is the refractory period following the stimulation Very beautiful reasoning demonstrates then, how the experimenting biologist may distinguish between the microphonic potentials and the neural potentials in spite of the fact that it is very difficult to prevent the experimenter's electrode from receiving each of these electric potentials with the other one Death interferes considerably later with the microphonic functions than with the neural functions.

The book does not regard the microphonic potential as the immediate cause of the neural potential, but makes it plausible that the microphonic potential creates a chemical which plays the role of a mediator between hair cell and nerve cell. The chemical in question must be one very quickly destroyed again. Examples of such chemicals are known and are mentioned

Pathological conditions, spontaneous or artificially produced, have been used for and against the microphonic theory of the hair cells. In any such case "obviously many factors, mechanical, biological, and electrical, are involved" The reviewer strongly indorses the last statement as a most needed one The complexity of these cases, experience with the literature shows, is too easily forgotten and replaced by a dogmatic simplicity which suits many another experimenter "Many experiments are inconclusive because of such complications," the book may be quoted.

The microphonic potential seems to be negative at the cilia of the hair cells, positive on their bottoms, that is, toward the basilar membrane How much its magnitude depends on the magnitude of the displacement of the membrane is a problem. How the displacement of any segment of the basilar membrane changes with increasing pressure, is totally unknown The enormous importance of the nature of the mathematical function of the two variables just mentioned, is pointed out on page 349. This critical attitude contrasts sharply with the dogmatism concerning the same problem which prevails in the first half of the book, where we read (p. 20) that "the ear does its analyzing as though it contained just such (*sic! Helmholtzian!*) resonators." How can such a statement be reconciled with "the case of the missing fundamental," already quoted by the reviewer? It is a flat contradiction. No Helmholtzian resonator records that which is missing, filtered out before it reaches the resonator.

The book then points out that the refractory period of each single nerve cell does not interfere with pitch conduction by the nervous bundle, because a nerve cell never comes into play singly, no matter how "placed" in the array. The book (p. 402) quotes Troland as having first suggested in print that the frequency is carried in a group of cooperating fibers. Wever and Bray called the same principle "the volley theory." The book criticizes the term "volley" as being inappropriate because it implies that all units fire at once; it proposes as a more apt military analogy "platoon fire." The larger group, the battalion of nerve cells, can by such platoon fire carry any microphonic frequency of the hair cells, without limit. The hair cells are true microphones, have no refractory period, so that the problem of frequency limit does not exist for them.

With respect to the higher levels of the nervous system the book believes that in the synapses there is no direct electrical conduction, but that a chemical is produced as a mediator, like the mediator between the hair cell and the nerve cell. The delay in a synapse is around one-half of a millisecond. This has some relation to the apparent direction of a source of sound. The synaptic latency is diminished with increasing stimulus intensity.

Two brief appendixes are added to the text. The first is merely a mathematical transformation of a formula for pitch modulation in physics. The second makes very interesting, even shocking, reading matter. We are now told that all belief in tuned elements (including those assumed in the cochlea by Wegel and Lane) may just as well be set aside. To quote:

> "It is also possible to treat the cochlea as a hydraulic system contained in a vessel with elastic walls.—Thus it is possible for the basilar membrane to act as an analyzer, and at the same time show no free vibrations after the stimulus has ceased."

And continuing: This was done by "Reboul in 1938." Reading this astonished the reviewer, because exactly it he treated elaborately ten years ago in his article "The Hydraulic Principles Governing the Function of the Cochlea" in the first volume of the *Journal of General Psychology;* and he had shown there that the basilar membrane would even "analyze out the missing fundamental," such a mystery otherwise. In a more popular, i.e., less mathematically analytical and instead more graphic way, indeed, he has published the same idea in many publications ever since the beginning of this century; he built his first hydraulic model in 1900 and published it in Pflueger's *Archiv fuer die Gesamte Physiologie.*

Naturally he looked into the bibliography at the end of the book, and he found there his name mentioned just once, and as the author of this hydraulic theory of 1928. But since neither his name as author, nor this article, nor his hydraulic theory of the cochlea, are mentioned anywhere in the text of the book, he wondered how this title got into the bibliography at the end. It must be through the mistake of a clerical helper. She probably copied titles from some other bibliography and innocently included "Max F. Meyer, The Hydraulic etc." not knowing that the authors intended to omit this item from their book. Whether his competitor Reboul in 1938 solved the hydraulic problem more exactly than he tried to in 1928, the reviewer cannot tell here, since he has not yet seen Reboul's article which was published in France. If Reboul thinks (as the authors hint without giving it a critical thought) that the cochlea and a blood vessel are hydraulically alike, he must be wrong. The tension conditions caused by the wall displacement are different.

A retrospect upon the book impresses one as follows. The authors when beginning the book started from a dogmatic conviction that the cochlea contained a statically stretched membrane consisting of an array of tuned resonators, à la Helmholtz. But during the progress of the manuscript their conviction became more and more shaky. Having almost finished they decided to go over the manuscript and add in convenient places where resonance is mentioned the little phrase "as if" meaning the same as "cross your fingers." Then when the manuscript seemed ready for the printer, Reboul struck them with his hydraulics, which gave them the knockout. Rewriting the whole book from a less dogmatic, more critical, point of view seemed too much of a task. So they saved their faces by just adding Appendix II

In this uncomfortable position seeking some further, even emotional, support, they fell back upon the unfortunate slogan introduced a dozen years ago by one of the Harvard professors to whom the book is dedicated, to wit, "place theory versus frequency theory" Because place and frequency are not absolutely and logically contradictory it is a slogan, stirring up dogmatic enthusiasm, but destroying a critical attitude Even microphones, by the thousands, must be placed somewhere, "spread out over a lamella not too stiff," as the French otologist Bonnier (not mentioned in the book) emphasized 50 years ago against the notion of Helmholtzian resonators Saying *Place versus Frequency* is like saying *Guelphs versus Ghibellines* or saying *Reds versus Royalists* At the end of their book the authors adopt the slogan and confess their belief in "a place theory without resonance," but nevertheless "a place theory of pitch perception"

But what good does all that beautiful exposition of the hair cells as piezoelectric microphones do, if in the end the hair cell is not a microphone but only "a place where" so to speak steel meets flint and starts a fire right on that candle and nowhere else and that is all?

During a half century about 57 physiological experiments have been made on the living animal with the preconceived purpose to "prove" the resonator theory, or call it place theory if you wish. The reviewer has often been asked if the 57 "evidences" did not overwhelm his obstructionism His reply is *No*, and for two reasons. (*a*) Each and everyone of the 57 is inconclusive, since it reveals among its results internal contradictions to one who has sharp enough eyes to see them, often they are declared by the experimenters themselves. (*b*) If one takes the standpoint of the hydraulic theory (which those experimenters and interpreters have never done because it would have been contrary to their purpose), he generally finds little difficulty in understanding why this place and that place and a further place of the cochlea should have "acted up" From seeing this riot to rejecting the hydraulic theory is a far cry. *"Place versus Frequency"* is a slogan unworthy of guiding scientific thought

And after all there is one argument of tremendous power (in the reviewer's mind) against any place theory of pitch perception A place theory is irreconcilable with an explanation of the esthetic effects of tones in melody and harmony Music is no perception of places This book, entitled "the psychology of hearing," abstains from this aspect of hearing

The book by Stevens and Davis is unequal in its parts. The latter (neurological) part is a masterpiece of original, penetrating, and critical thought and experimentation. Of the preceding part the reviewer cannot say the same. That is a compilation. But it is an assiduous compilation of the acoustic literature of the last six or eight years, and as such it is of high value if read with a critical eye. The whole book is one for which no person interested in the sense of hearing can afford to leave unreserved a place on his book shelf.

Coconut Grove
Florida

(*Lewin, Kurt. Principles of Topological Psychology Translated by Fritz and Grace Heider New York McGraw-Hill, 1936. Pp 231.*)

REVIEWED BY A S. HOUSEHOLDER*

In the *Principles of Topological Psychology*, Professor Lewin urges the employment by psychologists of a restricted mathematical idiom, that of the branch of geometry known as topology, for the purpose of formulating psychological "laws" with which "predictions" can be made in concrete cases. In this review we wish to try to summarize the author's program for the formulation of these laws in terms of this particular idiom and to evaluate the author's achievement in the light of his own explicit requirements To this discussion we wish to append some remarks concerning the role of any mathematical idiom in an empirical science and to relate these remarks to the author's employment of the idiom of topology in psychology It is no part of this reviewer's purpose, however, to concern himself in any way with questions of "truth," or of the ultimate validity of any laws that are or may be formulated in this idiom

True "constructive" concepts are required, according to this author, in terms of which "laws" can be formulated so as to make prediction possible We may say that the behavior of an individual is a function of the situation, $B = f(S)$. "In this equation the function f, or better its general form, represents what one ordinarily calls a law" (p. 11). The prediction is made in a specific case by substituting into the equation the constants which characterize this case But the "situation" must be taken as sufficiently inclusive, embracing all relevant factors In particular, one cannot hope to formulate valid psychological laws if one neglects the will, the needs, and the personality of the individual Better than "situation" is the expression "psychological life space," which is used "to indicate the totality of facts which determine the behavior of an individual at a certain moment" (p. 12).

Basic among the characteristics of the life space at a given moment are what is possible and what is not possible Hence the concepts must be so constructed as to reveal these features clearly Otherwise put, these facts are among the most important of those which determine the "structure" of the life space. But it must be understood that the "possible" or the "not possible" for the individual is not necessarily the objectively possible or not possible More generally, the structure of the objective physical, social, or conceptual environment is not necessarily identical with the structure of the psychological life space "These [the objective] facts are to be included in the representation of the psychological life space only to the extent and in the manner in which they affect the individual in his momentary state" (p 24). To carry out this distinction the terms "quasi-physical," "quasi-social," and "quasi-conceptual" are introduced for describing the corresponding facts in the life space

---

*Ph D in mathematics, University of Chicago General Education Board Fellowship in mathematical biophysics, University of Chicago

This distinction made, we learn that there is no need to speak of the future events as determining the present behavior, though the "time index of a psychological fact" may point to the future Again, while "mental events" are often indeterminate in greater or less degree, this indeterminateness is "an indeterminateness of the content and not of the psychological fact itself" (p 40). Time indices and degrees of indeterminateness are themselves psychological facts which are included as elements of the total structure of the life space

Then what is to be taken as "real" in the life space of the individual at any given time? Consciousness is not a criterion, for there may be many facts operating of which the individual is not at the moment conscious. "As far as the conceptual derivation is concerned, one may use effectiveness as the criterion for existence '*What is real is what has effects*' " (p. 19, author's italics)

Thus the purpose of the constructive concepts in psychology is to obtain laws and predictability The criterion for constructing the life space to be described by these concepts is effectiveness. Now what is to be the nature of these concepts? The answer is implicit in the terminology so far employed They are to be spatial and dynamical in character. "I am convinced that these concepts which we use for the representation of psychological facts, like region, spacial (*sic*) relationship in life space, connectedness and separateness, belongingness, etc., are real spacial concepts in a strict mathematical sense" (p. 42).

In passing we may note two peculiarities in the wording of this sentence. One is the assertion of conviction, which, in such a connection, is certainly a tacit admission of failure to demonstrate the proposition of which the author is "convinced." In fact—and this is the second peculiarity—it is by no means clear what is meant by saying that these concepts "are real spatial concepts in a strict mathematical sense" Presumably it means that the geometrical idiom can be conveniently employed in the building of a scientific theory of psychology There is only one way to demonstrate that this is so, and that is to build such a theory and exhibit it. Apparently this is the object of the book, and it is the purpose of this review to consider whether such a theory has been constructed. We return, therefore, to the book.

Within the total life space are to be distinguished regions separated by more or less pervious boundaries, or barriers One of these regions is the individual himself. The regions and their boundaries (together with their interrelations) represent quasi-social and quasi-conceptual facts as well as quasi-physical facts They are by no means static, but are always subject to change ("restructuring") as a result of outside influences from the environment or of processes within the individual himself To these changes they offer varying degrees of resistance They also exert mutual attractions and forces of various kinds, especially upon the individual himself.

We are told further that a life space may be more or less "structured," may have more or fewer distinguishable regions, but even where it is "infinitely structured" and subregions can be distinguished *ad infinitum* the concepts of distance and direction do not seem to apply (p 55) For this reason the concepts are "topological," since topology is that branch of geometry which has to do with the properties of space that are independent of the notion of distance It is true there do seem to be cases where it is meaningful to speak of greater and less distances, or of the shrinking of a region, but these can be shown on examination to be capable of purely topological characterizations like the rest (pp 156 ff).

Besides the generally limited structure of the psychological space there is another important difference between it and physical space. There is only one connected physical space, and from a particular physical situation S_1 a succeeding situation S_2 can always be derived uniquely, provided only the situation S_1 has been made sufficiently inclusive But the psychological space is subject to outside influences at any time, and when such influences are felt, not even the knowledge of all the psychobiological laws would make it possible to derive the situation S_2 from S_1 in psychological space Thus the physical world is dynamically closed, the psychological world not closed

Now, with Reichenbach, the author distinguishes two kinds of questions in the application of mathematics to empirical science, or, in this case psychology. These are the problems of pure mathematics, and the "problems of coordination" The latter problem is empirical, and its task is "to demonstrate the nature of certain empirical facts, in our case facts of psychological dynamics, and to coordinate to them mathematical concepts which represent the logical structure of these empirical facts adequately" (pp 59-60). It is desirable, so far as possible, to "choose coordinating definitions in such a way that they hold without exception and are univocal" (p. 93). Two types of mathematical tools are employed, topological and vector "In general one may say that the topological tools allow us to determine which events are possible in a given life space and which are not possible Vector concepts are necessary to determine further which of the possible events will actually occur in a given case" (p. 85)

The author's purpose, and the object of topological psychology, is abundantly clear and unambiguous to obtain psychological laws which yield predictions that are valid at least so long as no untoward disturbances are felt from the outside environment. These laws are to be expressed in terms of "constructive" concepts, spatial and dynamic in character, specifically designed so as to give an adequate representation of those aspects of the situation that are effective in the psychological processes. In order to apply these concepts and obtain these laws the coordinating definitions must be formulated, establishing a correspondence which is "without exception" and "univocal," between empirical facts and these concepts These definitions would thus comprise a kind of dictionary enabling one to translate from the language of direct observation into the mathematical conceptual language and back again

With such a purpose one can scarcely find fault. Furthermore, if usable coordinating definitions are given and some laws—or even one—are given by means of which one can actually make predictions that can be tested against experience, then one can feel only enthusiasm over the result. Such predictions as are made by a new theory might, indeed, fall wide of the mark, and this reviewer would not thereby condemn the theory But he would insist emphatically that if a theory is to be worthy of any serious consideration, it must yield predictions that are definite so that by testing against experience in specific cases it can be determined unambiguously (at least in principle, ruling out considerations of errors of observation, difficulties in isolating "pure" cases, etc) with what accuracy the predictions are fulfilled

Before returning to the author let us digress, for illustrations, to the field of physics The author himself invokes Reichenbach for the notion of coordinating definitions Moreover he opens his book with a discussion of the transition, in physics, from the "Aristotelian" to the "Galilean" concepts, a transition for which, he says, psychology is now ready, and which, in fact, it is undergoing It is quite legitimate, then, to note how a physical theory works

Not to complicate matters too much, we can say that a "situation" can be characterized by assigning to each point of a suitable region of space at the time considered the values of a certain number of parameters These parameters may be density of matter or mass of a material point, components of velocity vectors, components of force vectors, etc The parameters may or may not be independent, they may or may not be measurable directly by physical operations They may correspond to properties of purely postulated objects (e g, electrons) A "law," of the author's form, $B=f(S)$, describes the variation in time of one or a group of these parameters in terms of the present values of some set of them. That is, it defines a set of logical or mathematical operations which, if performed upon the present values of these parameters, and the value of the time-interval, yields the value of one of these parameters, or the values of several, after the lapse of this time-interval Hence if the law is to be applicable and a prediction is to be made, the necessary parameters must be obtainable by perfectly definite operations, whether of a physical or of a logical character, whether by means of empirical or merely postulated relations. But if this requirement is fulfilled, and if, further, a suitable "univocal" condition is fulfilled by the coordinating definitions, then from the calculated values of the parameters at the future time an interpretation is possible, by which at least some of the empirically observable characteristics of the future event can be described

Now we may return to topological psychology to see how it determines these parameters and makes its laws We first consider the necessary coordinating definitions, and then we shall consider their application to the investigation of a given situation.

For present purposes three of the coordinating definitions will suffice that of region, path, and boundary "Psychological region" is defined in this way "To each part of the life space a region is to be coordinated" This is to include "everything in which an object of the life space, for instance a person, has its place, in which it moves, . " "The reverse of the definition also holds, everything that is shown as a region in representing a situation must be a part of the life space" (p 93)

Next, "A path is to be coordinated mathematically to each psychological locomotion" (p 95) Conversely, "To each path in the psychological life space corresponds a locomotion which can or cannot be carried out" (p 96) The author seems a bit dubious about the "can or cannot" qualifier, so he goes on to explain

> Analogous definitions are common in mathematics Thus one usually defines function as follows *a* is called a function of *b* if the value of *a* varies, or does not vary, with the value of *b*.
>
> On the basis of these coordinating definitions of the psychological region and the psychological locomotion one can represent mathematically the topological relationships of an unlimited number of different situations It is the task of the single psychological investigations to carry out these constructions (p 96)

Thus the author believes that these two definitions alone suffice in many cases, though not all However, he adds a third, which exhausts the topological elements or objects, as opposed to their topological properties "We shall designate as the boundary of a psychological region those points of a region for which there is no surrounding that lies entirely within the region" (p 118)

Now let us proceed to try actually to "represent mathematically the topological relationships" of one of the "unlimited number of different situations" for which it is claimed that these coordinating definitions suffice, and so to determine the parameters in the equation $B=f(S)$ Unfortunately, before we can even take a further step we must go somewhat beyond the book itself and try to express more precisely notions that are only vaguely stated in the book. Of course, by so doing one runs the risk of making a misrepresentation, but the responsibility rests entirely upon the author for not having done it himself, since such precision is necessary before one can claim to have used topology in psychology

Let us try to strip away all possible complications. Let us suppose *a very simple case in which the individual is about to perform* a purely physical locomotion, with quasi-physical facts the only ones significantly contained in the structure of the life space Let us imagine that there is an individual and his goal, spatially separated, so that its attainment requires the individual's physical locomotion across territory unknown to him It seems reasonable to suppose, then, that the life space is two-dimensional, consisting of two-dimensional regions separated by one-dimensional (polygonal) boundaries Some of the space, we are told, may be unstructured, the rest structured Suppose the structured parts to consist of clearly defined regions, the indefiniteness being confined exclusively to the unstructured parts It seems that regions can exert influences of various sorts upon one another and upon the individual, but suppose the mutual influences to be completely represented by the topological character of the space (i e., which regions are adjacent to which, which are connected, etc), and those upon the individual to reduce solely to a constant attraction exerted by the goal The possible dynamical properties of the regions and boundaries are nowhere completely enumerated, but suppose (still for greater simplicity) that there is only one type of intrinsic quality significant to this case, viz , resistance to locomotion Let us say, then, that each region is characterized dynamically by a coefficient of friction, and each boundary segment by a permeability coefficient Suppose, finally, that even the unstructured region can be suitably characterized by assigning to it a suitable coefficient of friction which represents somehow the individual's estimate of the number and character of the possible barriers across it. Thus no distinction needs to be drawn, in this simplified case, between the structured and the unstructured regions

In this highly simplified case the necessary parameters may be summarized as follows To each region corresponds a coefficient of friction, to each boundary segment corresponds a permeability coefficient; to each region every boundary segment can be associated to form a pair, and to every such pair an "incidence-number" can be associated which has the value of *0* if the segment does not, *1* if it does bound the region. The situation is then completely described by a certain set of numbers

Now let us try to apply topological psychology to make a prediction as to the locomotion that will take place. Take first the simplest problem and suppose we know already the values of the necessary parameters What "law" do we apply? Not one is given A 200-page book on this allegedly "predictive" science of topological psychology states no laws—either as discovered, derived, or postulated.

Still, there does seem to be implicit in the course of the argument a kind of "least action" principle It might run somewhat as follows Among all paths in the life space joining the individual and his goal, he will take that path which offers the least total resistance Now if some relation between friction (of regions) and permeability (of barriers) is given so that we know how to combine the resistances along these possible paths, we should have a simple minimizing problem for finding the path to be taken It may be that such a law could be formulated. However, it is not formulated by the author. But even if it were formulated it would then be necessary to determine the necessary parameters that were at first supposed to have been known. And again we are left in the dark, or at least this reader has not been enlightened.

Suppose we go back to the coordinating definitions, say that of region. A region is anything that is "a part of the life space." How do we know what is a part of the life space? It is *not* what is a part of the physical space—at least, not necessarily, not until it has been shown to be so But how shall we show it? Consciousness is not a criterion, as we have been told, even in the case of an adult human being, and so a verbal report is at best insufficient It is true, we are given the definition of "reality" as that which has effects When has it the effects? Must we wait for the future effects, the ensuing behavior? Then where is the prediction? In terms of the future event we describe the present situation topologically in order to predict the future event! Or are present, contemporary effects meant? If so we are back to where we started. Just what effects do we look for, and what physical and conceptual operations must we perform to obtain the parameters describing what is "real"?

Perhaps we have not read far enough, so we return to the explanation following the coordinating definition. We find it explained that there are two ways of characterizing a region. Thus

> one can characterize a region by its qualitative properties and can find out *its relations of position* by determining which regions are contained in others (have the relation to each other of part and whole, $X \subset Y$), how they overlap ($X Y \neq O$ or $= O$), which regions have common boundaries (b), and which do not ($b_x b_y \neq O$ or $= O$)

Or alternatively, "One can build on psychological processes which connect different points (part regions) in the life space, for instance on locomotions" (p 94) As examples of "qualitative characterizations" are "ground which is easy or difficult to walk on, . ; a region of a certain color, the sphere of influence of a person; . . ." And again we seem to be back to where we started If the individual is not actually on the ground, how do we know what is the position and the walking difficulty of this ground as part of his life space? or, indeed, if it is a part at all?

Still hopeful, we look at some of the author's own concrete examples. Often such examples will help to clarify what is in part concealed in the short-hand of the technical jargon There are many such examples, but we shall choose two of the shorter ones, more or less at random These seem to be fairly typical.

> A mother has taken a year-old child away from play and wants to feed him on her lap He does not want to eat He is *at the moment dominated by the tendency "away from eating"* or "toward play" The mother holds the child on her lap and prevents the intended movement "away from eating" She puts her arm around him so that he cannot break away The mother's interference has in this case the character of a barrier (b) [*the author gives a diagram to make this clear*] between the region of eating (e) and that of play (pl) This barrier at the same time keeps the child C from pulling away from the spoon (sp) as it is brought near his mouth The child now begins to play on the mother's lap The mother tries to put an end . (p 47)
>
> A one-year-old child C stands behind a U-shaped bench B [*another diagram is given to explain this*] He wants to get the toy T on the other side of the bench but has not yet a sufficiently broad survey of the situation to carry out the necessary detour In this case also there is a barrier for bodily locomotion which makes it impossible for C to reach T Therefore according to our definitions, C and T, as far as this locomotion is concerned, do not belong to a connected region The

> barrier has to be represented in this case also as a Jordan curve in spite of the fact that it is not closed physically Vector psychology has to take up in detail the conceptually difficult problem of detour (p 146) [*At this point reference is made to two of the author's papers where presumably this is done*]

Even this leaves us no closer to a solution of our difficulty than we were before Indeed as far as the reviewer can determine from a careful study of the "definitions" given, the explanations and the illustrations, the term topology is used only to lend an air of profundity to some exceedingly simple diagrammatic illustrations of the type the author claims explicitly to renounce (p 76).

It is therefore the conclusion of this reviewer that "topological psychology" is not yet even the hope of a science. It formulates no laws as either empirically given or postulated, and it gives no prescription of the operations required for determining the parameters of the situations which are alleged to be adequate for the statement of laws And let it be recalled that our discussion has been carried on entirely in terms of "quasi-physical" facts, whereas these "constructive concepts" are supposed to be suitable also for describing and predicting quasi-social and quasi-conceptual facts and behavior. With this in mind the hopeless indefiniteness of this "theory" becomes all the more evident

It should be noted again that we are not raising the question of whether the theories are in agreement with experiments We are simply looking for a precisely stated law or a clear and usable definition, and we find none

Before concluding this review it may be of interest to cast a glance over the mathematical equipment which the author regards as necessary for his investigations This is easily done since the author interposes a short chapter in which he explains these concepts of topology for the benefit of the non-mathematical reader. For example, we learn that if A and B represent two regions, then "By $A + B$, the 'sum' of A and B, one understands the totality of all points which are included in either A or B By the 'intersection' of A and B (the part which is common to A and B) is meant the totality of points which are part of both A and B. One expresses intersection as $A B$" "If A is part of B then $A + B = B$ and $A B = A$" (p. 87). These are difficult matters and we cannot discuss them further We shall be content to list some of the concepts defined

They are "connected region" (the region "Germany" on the map of Europe is not connected because of the Polish corridor), "open" and "closed" regions, "limited" and "unlimited" regions (just what they sound like), "simply connected" (having no holes) and "multiply connected" regions; "Jordan curves" (closed and not self-intersecting), "boundary" (what it sounds like), "path" (arc of a Jordan curve), and a few others equally recondite A good example of the manner in which these are "employed" is given in the passage from p. 94, quoted above, where the author elaborates upon the "qualitative definition" of region So far as this reviewer can make out, there is no instance where a single theorem from topology has been employed, and the concepts that are introduced from it would not tax the understanding of a mediocre high school quarterback.

That mathematics can serve a purpose in psychological theory and experiment no longer requires argument A mathematical idiom is employed—when it is employed—in any empirical science because of its brevity, and because of the logical relations that have already been revealed by the works of pure mathematicians. Our author himself refers to the theories of Rashevsky and of Thurstone A glance at their printed works can leave no doubt that without mathematics their theories could scarcely have been stated in understandable *form, nor could the consequences of these theories have been* deduced as readily.

By contrast topological psychology, which claims to be a predictive mathematical theory, neither makes predictions nor uses mathematics and thus renders scant service to either mathematics or empirical science

Department of Psychology
University of Chicago
Chicago, Illinois

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i e, the most recently received books will be found at the end of the list)

MACMILLAN'S Modern Dictionary New York, Macmillan, 1938 Pp 1466

PHELPS, H A Contemporary Social Problems (Rev ed) New York Prentice-Hall, 1938 Pp 820

STRANG, RUTH An Introduction to Child Study (Rev ed) New York Macmillan, 1938 Pp 681

MURPHY, G, & LIKERT, R Public Opinion and the Individual New York Harper, 1938 Pp 316

STEVENS, S S, & DAVIS, H Hearing Its Psychology and Physiology New York Wiley, 1938 Pp 489

MUNN, N L A Laboratory Manual in General Experimental Psychology (Rev ed) New York Prentice-Hall, 1938 Pp 286

McLEAN, D Knowing Yourself and Others New York Holt, 1938 Pp 275

MEAD, G H The Philosophy of the Act Chicago Univ Chicago Press, 1938 Pp. 696

RUCH, F L, & WARREN, N Working with Psychology Chicago Scott, Foresman, 1938. Pp 214

MUNN, N L Psychological Development New York Houghton, Mifflin, 1938 Pp 582

VARNUM, W C Psychology in Everyday Life New York McGraw-Hill, 1938 Pp 444

O'CONNOR, Z C The Runaway Boy in the Correctional School New York Columbia Univ, 1938 Pp 78

WAIT, W. T The Science of Human Behavior New York Ronald Press, 1938 Pp 335

ELLIS, W D A Source Book of Gestalt Psychology New York Harcourt, Brace, 1938 Pp 403.

MANOIL, A La Psychologie Experimentale en Italie Paris Alcan, 1938 Pp 489

WOODWORTH, R S Experimental Psychology New York Holt, 1938 Pp 889.

STINCHFIELD, SARA M, & YOUNG, EDNA H Children with Delayed or Defective Speech Stanford University Stanford Univ Press, 1938 Pp 174

BLUEMEL, C S The Troubled Mind Baltimore Williams & Wilkins, 1938 Pp 520.

STROUD, J B Introduction to General Psychology. New York Prentice-Hall, 1938 Pp 681.

THOMAS, G My Mind a Kingdom New York Dutton, 1938 Pp 294

KATZ, D, & SCHANCK, R L Social Psychology New York Wiley, 1938 Pp 699

POWERS, F F, McCONNELL, T R, TROW, W C, & SKINNER, C E Psychology in Everyday Living. New York Heath, 1938 Pp 511

REID, A C Elements of Psychology. New York Prentice-Hall, 1938. Pp 409

EWING IRENE R, & EWING, A W G The Handicap of Deafness New York Longmans, Green, 1938 Pp 327

LAPIERE, R T Collective Behavior New York McGraw-Hill, 1938 Pp 577.

MOLEY, R Are We Movie-Made? New York Macy-Masius, 1938 Pp 64

HEINRICH, W Travaux du Laboratoire de Psychologie Experimentale de L'Universite de Cracovie Paris Alcan, 1938 Pp 239

Allen, G C Japan The Hungry Guest New York Dutton, 1938 Pp 261

Ruch, F L, & Warren, N Working with Psychology Chicago, Scott, Foresman, 1938, Pp 214

Bennett, C C An Inquiry into the Genesis of Poor Reading New York Bur Pub Teach Coll, Columbia Univ, 1938 Pp 139

Jensen, A S Psychology of Child Behavior New York Prentice-Hall, 1938 Pp 664

Imus, H A, Rothney, J W M, & Bear, R M An Evaluation of Visual Factors in Reading Hanover Dartmouth College, 1938 Pp. 144

Wallack, W M The Training of Prison Guards in the State of New York New York Bur Pub. Teach Coll, Columbia Univ, 1938 Pp. 396.

Toops, H A, & Haven, S E Psychology and the Motorist Columbus, Ohio Adams, 1938 Pp 265.

Harpower, M R The Psychologist at Work New York Harper, 1938 Pp 184.

Jones, F E The Defence of Democracy New York Dutton, 1938 Pp 352

Annual Report of the Board of Regents of the Smithsonian Institution Washington Smithsonian Institution, 1938 Pp 580

Alihan, M A Social Ecology New York Columbia Univ Press, 1938 Pp 267

Terman, L M. Psychological Factors in Marital Happiness New York McGraw-Hill, 1938 Pp 474

Boas, F, et al General Anthropology New York Heath, 1938 Pp 718

Tinker, M A, & Baker, K H Introduction to Methods in Experimental Psychology New York Appleton-Century, 1938 Pp 222

Landis, C, & Page, J D Modern Society and Mental Disease New York Farrar & Rinehart, 1938 Pp 190

Skinner, B F The Behavior of Organisms New York Appleton-Century, 1938 Pp 457

Thorpe, L P Psychological Foundations of Personality. New York McGraw-Hill, 1938 Pp 602

Patty, W L A Study of Mechanism in Education New York Bur Pub Teach Coll, Columbia Univ, 1938 Pp 183

McFarland, M B Relationships between Young Sisters as Revealed in Their Overt Responses. New York Bur Pub Teach Coll., Columbia Univ, 1938 Pp 230

Beck, S J Personality Structure in Schizophrenia New York Nerv & Men Dis Monog, 1938, No 63 Pp 88

Rudolfer, N da Silveira An Introduction to Educational Psychology Sao Paulo Companhia Editora Nacional, 1938 Pp 459

Jacobsen, E You Can Sleep Well New York McGraw-Hill, 1938 Pp 269

Landtman, G The Origin of the Inequality of the Social Classes Chicago: Univ Chicago Press, 1938 Pp 444

Simpson, R H A Study of Those Who Influence and of Those Who Are Influenced in Discussion New York, Bur Pub. Teach Coll, Columbia Univ, 1938 Pp 89

Bender, L A Visual Motor Gestalt Test and Its Clinical Use New York Amer Orthopsychiat Assoc, 1938 Pp 176

Kirkpatrick, C Nazi Germany Its Women and Family Life New York Bobbs-Merrill, 1938 Pp 353

Murray, H A Explorations in Personality New York Oxford Univ Press, 1938 Pp 761

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

June, [illegible]
Volume [illegible], Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

John E. Anderson
University of Minnesota
Charlotte Bühler
Universität, Wien
William H. Burnham
Clark University
Cyril Burt
University of London
Leonard Carmichael
Tufts College
Ed. Claparède
Université de Genève
Edmund S. Conklin
Indiana University
Sante De Sanctis
R. Università di Roma
Arnold Gesell
Yale University
William Healy
Judge Baker Foundation
Boston
Leta S. Hollingworth
Teachers College,
Columbia University
Buford Johnson
The Johns Hopkins
University
Harold E. Jones
University of California
Truman L. Kelley
Harvard University

Yoshihide Kubo
Hiroshima Normal
College
K. S. Lashley
Harvard University
A. R. Luria
Medico-biological
Institute, Moscow
Toshio Nogami
Kyoto Imperial
University
Henri Piéron
Université de Paris
Calvin P. Stone
Stanford University
Lewis M. Terman
Stanford University
Godfrey Thomson
University of Edinburgh
E. L. Thorndike
Teachers College,
Columbia University
C. J. Warden
Columbia University
John B. Watson
New York City
Helen Thompson
Woolley
Teachers College,
Columbia University

If this space should be unmarked, this is the regular library edition. But if this space is stamped with a designating seal, this is a special edition, sold under the restrictions of a biological copyright, and may not be resold for a period of five years from date of publication.

Copyright, [illegible], by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter, August 1, [illegible], at the post-office at Worcester, Mass., under the Act of March 3, 1879.

Reentered as second-class matter [illegible], at the post-office at Provincetown, Mass., under the Act of March 3, 1879.

The Journal of Psychology

EDITED BY

Carl Murchison

The Journal of Psychology was founded at the request of many psychologists throughout the world for the purpose of establishing a medium in which publication shall be immediate and in which other publication experiments may be carried out. Beginning with January, 1936, the following methods and forms will be in force:

(1). The size of each volume will be 400-600 pages, and the subscription price will be $7 per volume, two volumes per year.

(2). The format of each volume will be such that each subscriber may take down the assembled volume into its component articles.

(3). Except through the fault of the author, the lag in publication should not be more than three or four weeks.

(4). The author will receive 200 reprints of his paper, and the subscriber will receive the accumulated publications bound into half volume units. The second half of each volume will contain the title page and table of contents for that volume. Each half volume unit will be mailed to the subscribers at its stated date.

(5). Each author will be charged the actual manufacturing costs of any tabular or half-tone cuts.

(6). The publication charges will be as follows:

4	page signatures	$8
8	" "	16
12	" "	24
16	" "	32
20	" "	40
24	" "	48
28	" "	56
32	" "	64
36	" "	72
40	" "	80

These charges are tentative and will be adjusted as justified by the subscription list.

(7). The accumulated units will be mailed to subscribers quarterly: January, April, July, and October. The units may vary in size, in accordance with publication needs.

Beginning with January, 1937, eight hundred to twelve hundred pages (two volumes) annually. Per annum $14.00; per volume $7.00; per half volume unit $4.00. Complete sets at $7.00 per volume, plus transportation.

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

$7.00 per volume
Single numbers $4.00

QUARTERLY
Two volumes per year

June, 1939
Volume 54, Second Half

Founded by G. Stanley Hall in 1891

THE PEDAGOGICAL SEMINARY AND
JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

JUNE, 1939

(Manuscripts are printed in the order of final acceptance)

A new standardization of the Ferguson form boards . . . 265
By Louise Wood and Edythe Kumin

The elementary school mark . . . 285
By Clarence Carl Moore

A comparative study of the difficulty of the 100 addition combinations 295
By Lester R. Wheeler

Some relationships between stair climbing and prone progression . 313
By Louise Bates Ames

(OVER)

Copyright, 1939, by The Journal Press
Provincetown, Massachusetts

Entered as second-class matter, August 3, 1897, at the post-office at Worcester, Mass., under the act of March 3, 1879

Reentered as second-class matter May 11, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The intelligence of a negro high school population in a northern city 327
By Albert Sidney Beckham

Social relations and personality patterning in preschool children . 337
By Eugene L. Horowitz and Randolph B Smith

Season of birth and intelligence 353
By R. Pintner and G Forlano

An intelligent-question method of teaching and testing . . . 359
By Hulsey Cason

The perceiving and thinking of normal and subnormal adolescents and children on a simple drawing task 391
By Theodora M. Abel and Jane B Sill

Nationality preferences and stereotypes of colored children . . 403
By H Meltzer

Quantified pressure stimulation and the specificity and generality of response in fetal life 425
By Leonard Carmichael and M F. Smith

A test of general ability in Fiji 435
By Cecil W Mann

SHORT ARTICLES AND NOTES 455

Psychometric test results in two cases of precocious puberty (Arthur L Benton and Edward A. Hagmann) . 455

Worries of college freshmen (Ruth Lunger and James D Page) 457

Changes in variability in "digit-symbol substitution" performance measured at the beginning and at the end of practice (D G Ryans) 461

An observation of the changes in variability of high, medium, and low "intelligence" groups, measured at the beginning and at the end of "digit-symbol substitution" practice (D G Ryans) 467

BOOKS 471

CRITICAL REVIEWS OF RECENT BOOKS 475

B. F Skinner *The Behavior of Organisms* (Reviewed by Ernst Wolf) 475

C Spearman, *Psychology Down the Ages* (Reviewed by Herbert Moore) 481

BOOKS RECENTLY RECEIVED 487

The Journal of Genetic Psychology, 1939, 54, 265-284

A NEW STANDARDIZATION OF THE FERGUSON FORM BOARDS*

Judge Baker Guidance Center

LOUISE WOOD AND EDYTHE KUMIN[1]

A HISTORICAL

*In 1920 Ferguson (3) published a description of a new form*board series, which represented a real advance in the field of mental testing The development of a sequence of six formboards, progressing in difficulty by fairly even steps and used as a unit, was an innovation Such a series has several advantages for clinical use. Having a considerable range of difficulty, it is applicable and interesting to widely different age groups It is easily administered and not unduly time consuming. In spite of wide-spread recognition of its worth, the test has failed to prove as valuable a contribution as the material itself would warrant, largely due to the inadequacies and meagerness of the normative data. Work on the test has been published from time to time. There have been revisions of Ferguson's procedure and scoring scheme and some new norms have appeared. The present standardization is based in part on this earlier work and is a further attempt to make the test of more practical value

Ferguson selected the six boards in his series from 12 trial boards on the basis of the performance of 170 subjects in the elementary grades and college Then a scoring scheme was developed and grade norms computed on 364 elementary, high school, and college subjects including the original 170 A correlation coefficient of .81 was obtained between Ferguson score and grade placement Thirty-six pupils from the sixth grade were given the Army Alpha, and the correlation between this test and Ferguson score was 51 Teachers' estimates of intelligence and class standing gave correlations of 50 and .56 respectively with the formboards No significant sex difference was found

---

*Received in the Editorial Office on May 5, 1938

[1] The authors wish to express their appreciation to E. Morton Jellinek for his advice and assistance in the treatment of the statistical data of this study.

In 1927 Gillespie and Brotemarkle (5), in connection with the study of college student personnel problems, published percentile norms based on time for completion, for Boards IV, V, and VI separately, apparently treating each as an individual test Two years later they revised their norms, added numbers causing definite changes in their ratings It is significant that they found the upper half of the test discriminative at the college adult level. At about the time of their first work Plant (1) revised the directions and Shimberg (1) introduced a new scoring scheme, with tentative norms based on clinic cases MacPhee and Brown (9), in 1930, reported *An Inquiry into the Standardization of the Ferguson Form Boards.* The study was based on a small group of children aged eight to twelve years, selected by age regardless of grade. The original Ferguson scoring and directions were followed. The scores showed no significant increment at successive age levels. An alternate scoring scheme was devised which failed to give reliably discriminative values Only a small part of the group was retested six to eight weeks later, with a resultant high reliability coefficient Some sex difference in favor of boys was found.

In 1932 and 1933 two articles discussing the use of the test in immigration work appeared. Kolb (7) worked with several thousand subjects from six nationalities, while Reichard's study (12) was based on Polish people only Both were adult groups, 15 years and up. Kolb made comparisons in test performance between literates and illiterates, skilled and unskilled workers, men and women, by age and nationalities Again a sex difference in favor of men was seen Skilled workers excelled unskilled Reichard used both the Shimberg and Ferguson scoring plans for his group of 369 subjects, finding a correlation of 93 between the two. Both of these writers, using a battery of performance and other tests, expressed preference for the Ferguson Boards for their work with immigrant groups

Using first 1000, and later 500 inmates of a penitentiary, Pescor (10, 11) published in 1935 and 1936 two articles dealing with the Ferguson Form Boards and their relation to school achievement and language and non-language intelligence tests. The nationality and race of both groups were mixed. The average chronological age of the first group of subjects was 33 years, and while none were illiterate, about 60 per cent had attended grade school only. The average age of the second group was 38 years, and 65 per cent either had no education at all or had attended primary grades only. Using the Shimberg scoring with an arbitrary expansion of the range of mental age ratings, he compared it with the original Ferguson scheme, finding the latter preferable because it gave a better distribution and was more discriminative at the upper levels. His results showed that skilled workers gained higher scores than unskilled, and educational achievement apparently had little influence on performance The coefficient of correlation between the Ferguson Test as scored by the expanded Shimberg method and the Stanford-Binet for 1,000 subjects was .30, the correlation ratio being .35 Ferguson's scoring system gave about the same results. For the 500 cases, the correlation of the Ferguson Test and the Army Beta Test was 50 The correlation with Stanford Achievement Test for 219 cases from the same group was .15, although a slightly greater relationship was found in the earlier group

Very recently Williams and Lines (14) have published a restandardization of this test with age and grade norms based on a new scoring method They have radically revised the procedure in several respects, so that the character of the test itself seems to be somewhat changed. In presenting the blocks for each board, the directions call for a standard arrangement, which includes the turning of certain blocks as suggested earlier by Plant. All six boards are given regardless of failures, and failed boards are completed in front of the subject. The emphasis is thus placed on learning rather than on the subject's ability to deal with a series of concrete problems

B SCORING

Ferguson based his scoring scheme entirely on the time taken for completing each board The values ranged from one to five points for each of the six boards alike A time limit of five minutes was set for each board The total number of points for the six boards gave the final score (Table 1).

TABLE 1
FERGUSON SCORING

0 - 29"	5 points
30"- 59"	4 points
60"- 99"	3 points
100"-149"	2 points
150"-300"	1 point
Failed in 300"	0 point

The Shimberg scoring plan published in 1927 provided a range in values of one to ten points for each board with the time intervals lengthened in accordance with the increasing difficulty of the boards. It did not take into account blocks correctly placed at the expiration of the five-minute time limit, but it was noted in the publication that such a procedure would probably augment the discriminative values of the scores

The scoring system used as the basis for the present work is a further revision by Shimberg.[2] The five-minute period was divided into fifteen 20-second intervals A value for each interval worked out for each board separately was then reached through statistical procedure In addition, values were computed for boards partially solved at the end of 300 seconds. In comparison with the earlier scoring plans, we have found this one to be distinctly superior

One of the chief criticisms of the first Shimberg scoring plan, as well as of the original Ferguson scheme, was the failure to give adequate credit for partial completion of the more difficult boards It is obvious that the range of difficulty in the series is such that partial completion of the last boards involves as much or more ability as is needed to complete the easiest boards within the time limit. Therefore, assigning a score of zero to all failed boards would, as several others have suggested fore-shorten the distribution and conceal variations apparent in the performance In developing the revised plan, Shimberg analyzed statistically the frequency of placements in partially completed boards The number of blocks correctly placed in the failed boards was found to be as discriminative as the time differences for the completed ones (Table 2)

The advantages of this new method of scoring are more concretely demonstrated by a comparison of the three scoring plans The performance of 100 boys taken from our group of subjects, 10 being selected at random at each chronological age level, was scored by all three schemes, the Ferguson, the original Shimberg, and the revised Shimberg plans. The inter-correlation of these results by the Pearson formula did not, as was to be expected, reveal the differentiation (The Ferguson and the original Shimberg scores gave an r of 94, the original and revised Shimberg schemes an r of .96, and the Ferguson and revised Shimberg plans came between with an r of 95)

---

[2] An unpublished study based on data collected at the Judge Baker Guidance Center

TABLE 2
REVISED SHIMBERG SCORING PLAN

Time	Credits for board I	II	III	IV	V	VI
			Completed boards			
1- 20	7	10	10	10	10	10
21- 40	5	8	9	10	10	10
41- 60	3	6	7	9	9	9
61- 80	2	5	6	8	9	9
81-100	1	4	5	7	9	9
101-120	1	3	4	6	8	8
121-140	0	3	4	5	7	8
141-160	0	3	3	5	6	7
161-180	0	2	3	4	6	7
181-200	0	2	3	4	5	7
201-220	0	2	2	4	5	7
221-240	0	1	2	3	5	6
241-260	0	1	2	3	4	6
261-280	0	1	2	3	4	6
281-300	0	1	2	3	4	6
			Incompleted boards			
No blocks correctly placed						
10 or 11	0	0	1	2	3	5
8 or 9	0	0	1	2	3	5
6 or 7	0	0	1	2	3	4
4 or 5	0	0	0	1	2	3
2 or 3	0	0	0	1	1	2
0 or 1	0	0	0	0	0	0

Self-correlations, however, in terms of intra-class correlation revealed a definite statistical difference in the internal consistency of these three scoring methods The scores of the even numbered boards were added together and the sum of these was correlated with the sum of the scores of the odd numbered boards by means of *the intra-class correlation technique*[3] The Ferguson scoring method gave the extremely low intra-class correlation of .13, the first Shimberg scheme an intra-class correlation of .60, while the revised method yielded the substantial coefficient of 75

Aside from this statistical analysis, certain other differences are obvious. Of our 100 subjects, five achieved the maximum score of 60 by the first Shimberg scoring. By the revised plan only one individual scored as high as 49, as against a possible maximum score of 57. Furthermore, among all the cases used for this study none reached the maximum and only one scored as high as 55. Pescor also discarded the original Shimberg method because too many perfect scores were attained. It is clear, therefore, that the range of possible scores is distinctly more adequate in the revised than in the original Shimberg plan. In contrast, while no one of our cases, nor any of those reported by others, reached the maximum score according to the original Ferguson scoring, another criticism can be made here, in that a comparatively small proportion of the possible range is used Beside the results from the small group of 100 cases, in our long experience with this scheme, earlier used exclusively in the clinic, no record of a greater score than 22 is found This corresponds with Reichard's findings in his use of the test with adults 15 years of age and over, in which the maximum score among 369 subjects was the same Ferguson reports several scores above this, but only at the college adult level and none above 25 points. Through the twelfth grade, which roughly corresponds to our upper age limit of 17 years and 5 months, he found no higher score than 22 More concretely, scores do not cover 75 per cent of the total range offered by this scoring method, while over 90 per cent of the possible range is used by the revised plan The explanation lies in the construction of Ferguson's scheme The arbitrary assignment of a maximum 5-point value for any board completed in less than 30 seconds makes it practically impossible for an individual unfamiliar with the test to gain this credit on either of the last two boards Among our 100 boys the shortest time for completing the sixth board was 69 seconds. According to Ferguson's scoring this would be assigned a value of 3, while by the revised scoring a value of 9 points would be gained. Naturally, neglecting to give credit for partially completed boards also contributes to this foreshortening, especially as the final board is very frequently failed. It is noteworthy that of the 100 boys, 73 per cent were unable to complete the sixth board within the time limit, while 29 per cent failed the fifth

---

[3] The formula used was Jellinek's (6) transformation of Fisher's formula (4, pp 198, ff.)

C. PROCEDURE

The original directions for administering the test, as published by Ferguson, called for a haphazard removal of the blocks before the subject The procedure was later modified by Plant (1) who established a more standard method of presentation. This changed the problems presented by the individual boards and controlled, to some extent, the element of chance placements. We have taken over this major development, using Plant's directions without change.

The boards are presented in order until the subject has failed two in succession. Each is placed before the subject in the designated position. The blocks are removed out of view and placed outside the board, to the right for right-handed subjects, and to the left for left-handed subjects, in random fashion except that no two belonging in the same inset are adjacent. In each instance the examiner says, "Put these pieces in as quickly as you can." When the second board is presented, we have added the explanation that there are two blocks for each space instead of one. It is usually *unnecessary to repeat this for the rest of the set.* No account is taken of errors. The time for completion of each board is recorded. In case of failure to complete in the five-minute limit, the number of blocks correctly placed at the end of that time is noted. The total of the points obtained on the separate boards gives the final score.

D. Selection of Group

In experimenting with the boards it became evident that the test is too difficult for the average child of seven years or under. Consequently we included in our standardization children ranging in age from 7 years and 6 months to 17 years and 5 months. The total group was comprised of 3,012 children, 2,731 or 91 per cent of whom were seen at the Judge Baker Guidance Center. The remaining nine per cent were public school children who were included in order to increase the very small numbers obtained at the

TABLE 3

Age	Boys No. clinic subjects	 No. school subjects	 Total	Girls No. clinic subjects	 No. school subjects	 Total
7/6- 8/5	32	46	78	10	50	60
8/6- 9/5	97	14	111	21	46	67
9/6-10/5	126	2	128	45	52	97
10/6-11/5	184	1	185	50	46	96
11/6-12/5	201	0	201	85	21	106
12/6-13/5	201	0	201	104	3	107
13/6-14/5	200	0	200	165	0	165
14/6-15/5	212	0	212	200	0	200
15/6-16/5	200	0	200	200	0	200
16/6-17/5	203	0	203	195	0	195
	1,656	63	1,719	1,075	218	1,293

clinic on the lower age levels Table 3 gives the distribution of the entire group by age and sex, as well as the proportion of school subjects.

From inspection of Table 3, it will be seen that from 10 years 6 months up the numbers for boys approximate 200 at each age level and fall below 100 only on the lowest age level The number of girls falls considerably short of this, although the four upper levels compare fairly well with the boys' groups, and it is only on the two lowest levels for girls that the number does not approximate 100 at least. Percentile ratings were first computed for each of the upper age levels when only 100 cases had been obtained A comparison of these figures with the final ratings showed very small differences, never over one point and largely under half a point

The Judge Baker group was selected at random from the clinic intake, except for the omission of children having an *IQ* on the Stanford-Binet of 80 or below and all cases where the diagnosis indicated any sort of abnormality such as epilepsy, head injury, beginning psychosis, or psychopathic personality, or where there was any question of the reliability of the performance Children seen at the clinic are drawn from all walks of life and include many educational as well as behavior problems. The proportion of cases classed as delinquents is about one-third of the total intake The clinic records for the last several years show that of the families studied approximately one-quarter of the fathers are engaged in professional or white collar work, while somewhat more than a quarter are skilled workers, and the rest, or about 40 per cent, are unskilled workers. After dropping out all defectives, to correspond with the group used for standardization, the median *IQ* for children seen over the same years ranges from 95 4 to 103.9 and averages 98 4

The school group was secured from several schools in a suburban community giving a fairly representative cross-section of economic strata No children more than two years retarded in grade for age were included In many instances an *IQ* on the Kuhlman-Anderson Scale was available at the school In cases where the *IQ* was unknown and there was any degree of school retardation, or the child appeared dull, the teacher's estimate of his ability was obtained We feel reasonably sure that children having an *IQ* below 80 were ruled out by this method.

E Results

The raw scores gained from the revised Shimberg scoring plan showed good age differentiation and range The frequency distribution of these scores for each age and sex is given in Table 4

TABLE 4
Frequency Distribution of Scores

Age	7-6—8-5	8-6—9-5	9-6—10-5	10-6—11-5	11-6—12-5	12-6—13-5	13-6—14-5	14-6—15-5	15-6—16-5	16-6—17-5
					Boys					
Score										
55-59										
50-54							1	2	7	7
45-49			1	5	4	8	14	22	19	32
40-44		5	4	14	24	24	34	31	51	42
35-39	1	7	9	28	33	47	45	60	48	45
30-34	13	16	27	38	48	55	52	50	43	42
25-29	19	21	30	43	45	26	25	28	18	16
20-24	14	26	29	30	20	24	13	12	7	9
15-19	12	13	16	14	15	9	11	1	3	7
10-14	8	12	8	9	10	6	3	4	1	3
5- 9	9	9	3	3	0	1	2	1	2	0
0- 4	2	2	1	1	2	1	0	1	1	0
	78	111	128	185	201	201	200	212	200	203
					Girls					
55-59							1			
50-54							0		1	2
45-49				2	2	1	4	6	8	9
40-44		1	5	4	12	8	16	19	23	34
35-39	1	5	2	14	10	25	31	49	43	47
30-34	4	9	21	16	18	15	32	43	51	41
25-29	12	7	23	23	25	22	38	41	38	37
20-24	14	14	17	16	12	9	25	19	17	14
15-19	12	8	15	8	13	11	14	13	15	4
10-14	7	12	6	7	4	9	3	7	2	5
5- 9	9	6	5	2	6	4	1	3	2	2
0- 4	1	5	3	4	4	3	0	0	0	0
	60	67	97	96	106	107	165	200	200	195

From these distributions percentiles were computed for each age level for boys and girls separately The resulting curves indicate a fairly constant age progression. As the general trend is so consistent, it may be assumed that the irregularities are due to errors of sampling Consequently the curves have been smoothed in the belief that the resulting figures are more truly representative Figures 1 and 2 give the original percentile curves with the smoothing

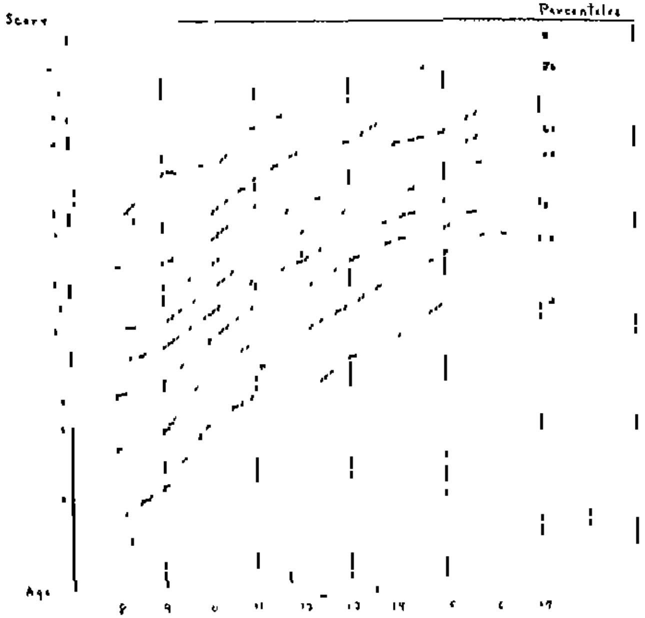


FIGURE 1
Boys

It is obvious from these figures that the curves for girls show many more irregularities than those of boys Only six slight corrections were made on the boys' curves, five of them for less than one point. On the other hand, 25 corrections were made on the girls' curves Twelve of these were under one point, and nine between one and two points, with none more than three points Tables 5 and 6 give the smoothed percentiles.

It will be noted that the girls' scores run consistently lower than the boys' at each age, and the difference tends to increase with age A comparison of the curves from plotting the median scores for boys and girls respectively shows this difference clearly (Figure 3)

A more detailed analysis of the data gives further evidence of

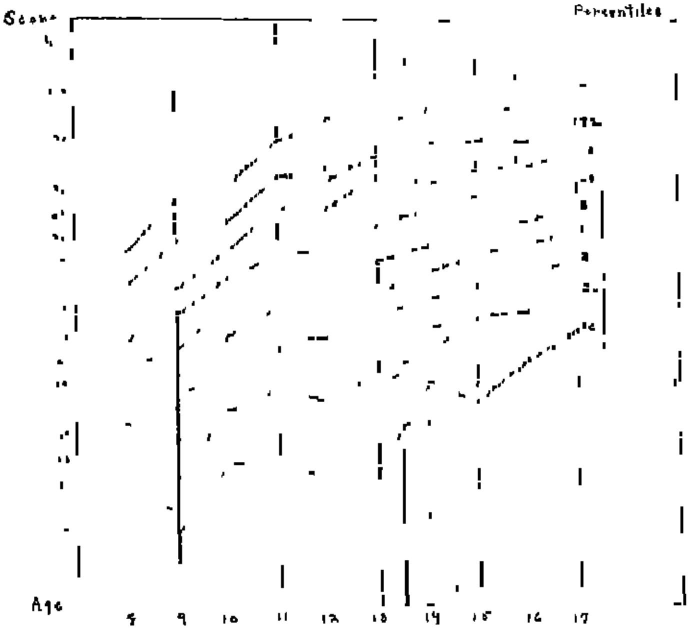


FIGURE 2
GIRLS

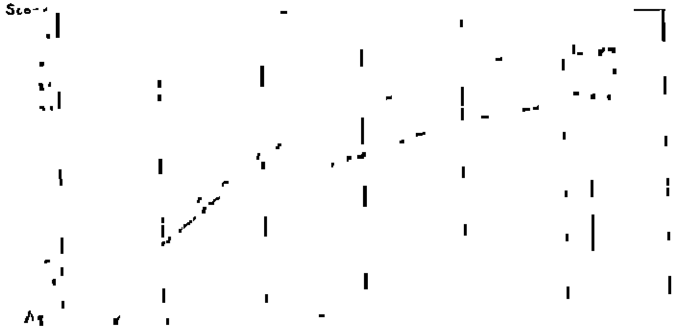


FIGURE 3
MEDIAN SCORES

TABLE 5

SMOOTHED PERCENTILES FOR BOYS

Age	7-6—8-5	8-6—9-5	9-6—10-5	10-6—11-5	11-6—12-5
Percentile					
90	32.1	35.9	37.1	40.3	41.8
80	30	31.4	32.7	35.9	38
70	28.5	28.6	30.5	33.5	35.5
60	25.4	25.9	28.3	31.3	33.6
50	23.1	23.7	26.3	29.2	31.2
40	20.2	21.5	24.5	27.3	28.8
30	17.6	18.6	22.3	25.1	26.4
20	13	14.7	19.7	22.2	23
10	7.4	9.7	15.3	17.4	17.7

	12-6—13-5	13-6—14-5	14-6—15-5	15-6—16-5	16-6—17-5
90	42.2	44.2	45.6	46.1	47.3
80	38.7	41.2	41.7	43.7	44.8
70	36.9	38.7	39.5	41.8	41.8
60	35	36.6	37.7	39.6	39.7
50	33.1	34.4	36	37.4	37.7
40	31.3	32.4	33.9	35.4	35.6
30	28.9	30.6	31.7	32.9	33.3
20	25	27.4	29.3	30.4	30.7
10	21	22.5	24.8	25.7	25.7

TABLE 6

SMOOTHED PERCENTILES FOR GIRLS

Age	7-6—8-5	8-6—9-5	9-6—10-5	10-6—11-5	11-6—12-5
Percentile					
90	29.6	33.9	35	39.2	40
80	27	30.7	32.3	36	36.5
70	25	26.7	29.8	33.3	34.2
60	22.5	24.7	27.7	29.5	29.9
50	20.3	21.8	25.4	27.2	27.6
40	18.8	19.6	23	25.2	25.6
30	15.8	16.7	20.3	22.6	22.7
20	12	13.3	15.5	17.7	18
10	7.7	8.9	10.5	12	11.8

	12-6—13-5	13-6—14-5	14-6—15-5	15-6—16-5	16-6—17-5
90	40.8	41	41.6	41.9	43.4
80	37.8	38.8	39.2	39.6	40.6
70	34.9	36.1	36.7	37	38.3
60	31.9	33.2	34.4	34.7	36.1
50	29	30.3	32.1	32.4	33.8
40	26.7	28.2	29.6	30.6	31.7
30	24.5	26.1	27	27.9	29.5
20	19.7	22	24.5	25.1	27
10	14.1	16.3	19.3	20.8	23.8

TABLE 7

Age	Boys N	Boys Min score	Boys Max score	Boys Mean± *SE*	Boys *SD*	Girls N	Girls Min score	Girls Max score	Girls Mean ± *SE*	Girls *SD*
7-6— 8-5	78	1	36	21.0±1.01	8.96	60	2	36	19.4±1.06	8.18
8-6— 9-5	111	1	41	22.7± .89	9.33	67	0	41	20.1±1.26	10.28
9-6—10-5	128	4	45	25.4± .71	8.07	97	3	44	24.1± .93	9.16
10-6—11-5	185	5	46	28.5± .63	8.53	96	1	47	26.2±1.04	10.16
11-6—12-5	201	0	48	29.9± .63	8.91	106	1	47	26.6±1.02	10.52
12-6—13-5	201	1	47	31.6± .58	8.27	107	0	48	27.4±1.02	10.53
13-6—14-5	200	6	51	33.4± .61	8.56	165	6	55	30.2± .65	8.41
14-6—15-5	212	3	52	34.8± .55	8.01	200	6	47	28.3± .58	8.23
15-6—16-5	200	5	51	36.3± .58	8.19	200	5	50	31.6± .58	8.17
16-6—17-5	203	10	53	36.4± .60	8.51	195	6	50	32.8± .59	8.30
Total group	1,719	0	53	31.3± .23	9.65	1,293	0	55	28.6± .27	9.86

the consistent superiority of the boys. Table 7 gives the minimum and maximum scores, the means with their standard errors, and the standard deviations of the distributions for each age. For the total group of boys the mean is 2.7 points higher than that for the total group of girls. This difference is 7.7 times as great as its standard error of .35, indicating a highly significant differentiation between boys and girls. The mean scores for boys are higher at every age level, the smallest difference being 1.3 and the largest 6.5. The small standard deviations of the means suggest a gratifying degree of reliability in these figures. Both the standard deviations of the average and those of the distributions tend to be somewhat larger for girls at the younger ages.

Since Kuhlman-Anderson *IQ*s were available for a number of the school children included in the group, a correlation between this test and the Ferguson Boards was made. One hundred forty-nine girls ranging in age from 7/6 to 13/5 and 44 boys between 7/6 and 11/5 had been tested on the Kuhlman-Anderson within a year of the time the formboards were given. An estimated mental age was computed on the basis of the Kuhlman-Anderson *IQ* and the chronological age at the time of administering the formboards. The Pearson coefficient of correlation for the boys is .47 and for the girls .45. The etas computed from the same figures came out slightly higher, .52 for boys and .54 for girls.

Naturally, larger numbers for correlation could be obtained from the clinic group. The subjects were divided into two sections, those having an *IQ* on the Stanford-Binet of 80 through 110, and those with *IQ*s over 110. For comparison a group of scores from clinic children rating below 80 *IQ* was added. These scores, of course, were not included in the standardization. The correlations between Stanford-Binet mental age and Ferguson score, and between chronological age and Ferguson score, for each of these intelligence levels, for boys and girls separately, appear in Table 8.

In the correlations of Ferguson scores with Stanford-Binet mental age the figures for the average and superior groups show a substantial relationship. For the group with *IQ* below 80 the correlations are considerably lower. The etas throughout are very little higher than the *r*'s, indicating, as did the scatter diagrams, that the relationship is linear. These results show no significant differences between the sexes. The correlations with chronological age for the same groups

TABLE 8

CORRELATION BETWEEN MENTAL AGE, CHRONOLOGICAL AGE AND FERGUSON SCORE

	IQ	*N*	*MA* and Ferguson r	*CA* and Ferguson r
Boys	Above 110	160	55	52
	80 through 110	243	54	53
	Under 80	202	36	25
Girls	Above 110	121	42	40
	80 through 110	161	53	52
	Under 80	180	.32	.24

are very similar except that the figures for the retarded group show even less relationship than with mental age. It is obvious from the similarity of the correlation coefficients for the average and superior groups that there is no statistically reliable difference between them. The correlations for the retarded group, being consistently lower, were compared with those for the other two groups by Fisher's z transformation of r (4, p 183, ff.). Table 9 gives these figures.

TABLE 9

z TRANSFORMATIONS OF r

	Boys				Girls			
IQ	r	z	*SE*	t_Δ	r	z	*SE*	t_Δ
			Stanford-Binet Mental Age and Ferguson Score					
110+	55	6184			42	4477		
—80	36	3769			.32	3316		
Δ	19	2415	± 1067	2 26	.10	.1161	± 1188	.9773
80-110	.54	.6042			.53	5901		
—80	36	3769			.32	3316		
Δ	.18	2273	± 0959	2 37	21	2585	± 1094	2 36
			Chronological Age and Ferguson Score					
110+	52	.5763			40	4236		
—80	.25	2554			24	2448		
Δ	27	3209	± 1067	3.01	.16	1788	± 1188	1 51
80-110	53	5901			52	5736		
—80	25	2554			.24	2448		
Δ	28	3347	± 0959	3 49	28	.3315	± 1094	3.03

For the boys, the ratios between the differences of the z transformations of r's and the standard errors of these differences show that the retarded group is significantly different from both the average and superior groups That is, all the differences are more than twice as great as their standard errors This holds true for Ferguson correlations with both chronological age and mental age For the girls, the differences between the average and retarded groups are also significant. The figures for the superior and retarded groups do not show a reliable differentiation, probably because the number of cases in the superior group was smaller On the whole, there is sufficient evidence that the retarded group is distinctly different from the others in Ferguson performance Consequently, any attempt to combine the three groups for purposes of correlation would give a meaningless coefficient

When the superior and average groups were combined, the correlation between Ferguson score and mental age was .54 for boys and 55 for girls, while the corresponding coefficients for chronological age were 47 and .33. A comparison with Table 8 shows that these figures for chronological age are relatively lower than those gained when the groups were subdivided, while the *r*'s for mental age show a very close correspondence except in the case of superior girls When partial correlations were computed, the correlations between mental age and Ferguson with chronological age held constant became .34 for boys and 47 for girls When mental age was held constant, the *r*'s between chronological age and Ferguson dropped to 17 for boys and .04 for girls.

F. DISCUSSION

The availability of clinic cases made it possible to obtain unusually large numbers for this standardization. With all abnormals ruled out and with all cases discarded where maximum effort and attention were questionable, the make-up of the resulting group should not affect performance on a test of this nature to any extent The clinic clientele is drawn from all parts of greater Boston and outlying districts and the socio-economic status of the clinic families is sufficiently varied to prevent any great sampling error of this sort. It is difficult to estimate the effect of the admixture of school subjects on the lower levels Unfortunately there are insufficient numbers of clinic scores at those levels to make an adequate comparison. It is apparent from an inspection of the percentile curves for girls that some consistent sampling error is present on the nine-year level. Aside from this, the irregularities are not systematic and show no distinction between entire clinic and the mixed school and clinic age levels. The curves for boys are quite regular throughout. Likewise the standard deviations for each of the age levels both for boys and girls show no consistent variations related to the mixed levels On the whole, it seems doubtful that there is any appreciable difference between the two groups

While there seems to be no question from the work of the majority of experimenters that the test itself is discriminative, it has already been noted that the degree of differentiation does seem to be influenced by the type of scoring scheme used McPhee and Brown (9) reported the one study which failed to show consistent age-increments by Ferguson's scoring or by a scheme of their own based on standard deviation intervals. Quite possibly this was due to the fact that their group was selected by age regardless of grade placement, with the result that many educationally retarded children were included Reichard (12) found, as we have, a very high coefficient of correlation between the Ferguson and the first Shimberg plans. He concluded, as a result, that a complicated mathematical treatment added very little to the discriminative value of the test, and therefore adopted the simpler Ferguson scoring method. As we have already shown by our comparison of the three plans, a further analysis reveals definite drawbacks not otherwise evident When an evaluation of internal consistency is made, the results of a scoring scheme developed by a statistical rather than an arbitrary treatment show to advantage In the light of this, the first Shimberg plan, in spite of its other defects, is far superior to the Ferguson The adopted plan not only has the highest internal consistency, but overcomes the disadvantages patent in the first Shimberg scheme. Because the test itself is discriminative, a high correlation between various scoring schemes is natural and therefore fails to appraise the individual methods.

A definite sex difference with males excelling females, both at school and adult ages, had been found by other experimenters with this test Our sex groups are not controlled for numbers, intelligence, or reason for referral to clinic. However, our results show such a consistent sex variation in favor of boys that irrespective of the statistical significance of the difference at any single age level, the total sex difference is significant Not only are both the mean and the median scores at each level higher for boys, but with one very slight exception on the lowest age level all the percentile scores are higher for boys than for girls of the same age Moreover, this difference increases with age As a result, the curve from the median scores for girls shows a greater flattening at the upper age levels, giving a different developmental picture from the boys Whether or not the girls are actually more variable in performance, as would be indicated by the fact that their standard deviations tend to be greater, is open to question It is possible that, since the groups were not paired, this may be accounted for by an error in sampling In general, however, in studies on performance tests, this type of sex difference has also been found.

Individual differences between subjects of the same age and sex are far greater than any sex or age differences which appear The range in scores at any one age level greatly exceeds the range of the median scores of the entire group, a span of 10 years in chronological age. For clinical work it is this individual variation which is chiefly important While the problem of the interpretation of the Ferguson Formboard scores has not been made an essential part of this study, some light has been thrown on the matter by statistics available here and in the work of others.

It is apparent from all the results that while intelligence is a factor in formboard performance, some other ability or combination of abilities is considerably involved The majority of investigators have found an *r* of about 50 between Ferguson performance and intelligence test results. Our correlations show striking agreement, except that a significant difference was found in the retarded group, where the test scores gave definitely lower coefficients. Consistent with this, it is interesting to note that where others have found low correlations between formboard and intelligence test scores, it is evident that their groups include a substantial proportion of the mentally retarded Between the normal and superior groups no significant difference in the correlations with mental age appeared It seems logical to assume that among retarded children intelligence is less of a factor in performance than a special aptitude.

In general, the results on the relation of chronological age to test performance are not entirely consistent or clear cut Certainly the test scores of the standardizing group show definite and steady age progression Of the smaller groups used for correlation, both the average and superior subgroups gave coefficients for chronological age corresponding closely to those for mental age and test score.

When these two groups were combined, the *r*'s for chronological age and test score dropped somewhat and partial correlations gave figures indicating a negligible relationship. In the defective groups, where the Pearson coefficients were much lower throughout, the same trend toward a relatively lower correlation with chronological age was found. Perhaps the seeming contradiction in these results was produced by the division of cases into groups according to *IQ* which automatically controls the relation between chronological and mental age. The total picture, however, suggests that the Ferguson scores are more closely related to mental than physical age development.

Some work has been published on the reliability of the test. Ferguson found a reliability coefficient of .80, but the number of cases and time before retest are not known. MacPhee and Brown gained a correlation of .90 on a very small number of cases with quite a brief interval between test and retest. We have some evidence that the interval between testings on these boards should not be too short, in order to prevent the influence of practice. An adequate investigation of the reliability of the Ferguson Test would be most desirable. The question of validity has not been touched in this study. Since it is evident that some special factor is involved in performance on the formboards, it remains to determine what this is and how far the test is prognostic for guidance purposes.

References

1. Bronner, A. F., Healy, W., *et al.* A Manual of Individual Tests and Testing. New York: Little Brown, 1927. (See pp. 126-127, 220.)
2. Brotemarkle, R. A. Individual mental testing at the college adult level. *J. Appl. Psychol.*, 1927, **11**, 415-436.
3. Ferguson, G. O. A series of formboards. *J. Exper. Psychol.*, 1920, **2**, 47-58.
4. Fisher, R. A. Statistical Methods for Research Workers. (5th Ed.) London: Oliver & Boyd, 1934.
5. Gillespie, F. R., & Brotemarkle, R. A. Interpolated revision of college adult level norms for personnel procedure. *J. Appl. Psychol.*, 1929, **13**, 325-345.
6. Jellinek, E. M. On the use of the intra-class correlation coefficient in the testing of the difference of two variance ratios. (To be published.)
7. Kolb, L. The intelligence of immigrants as measured by tests. *Proc. Amer. Assoc. Stud. Feeb. Mind.*, May, 1932.
8. MacFarlane, M. A study of practical ability. *Brit. J. Psychol., Monog. Suppl.*, No. 8, 1925.

9. MacPhee, E. D., & Brown, A. J. An inquiry into the standardization of the Ferguson form boards. *J. Educ. Psychol.*, 1930, **21**, 24-36.

10. Pescor, M. J. A statistical study of the Ferguson form board test. *Pub. Health Rep.*, 1935, **50**, 1815-1823.

11. ———. A further study of the Ferguson form board test. *Pub. Health Rep.*, 1936, **51**, 1195-1201.

12. Reichard, J. D. The intelligence of the prospective immigrant. *Pub. Health Bull.*, 1933, **206**, 10-11.

13. Schieffelin, B., & Schwesinger, G. Mental Tests and Heredity. London: Galton, 1930. (See pp. 166-167.)

14. Williams, G. W., & Lines, J. An evaluation of the Ferguson form boards and the derivation of new age and grade norms. *J. Appl. Psychol.*, 1937, **21**, 556-571, 673-687.

Judge Baker Guidance Center
38½ Beacon Street
Boston, Massachusetts

The Journal of Genetic Psychology, 1939, 54, 285-294

THE ELEMENTARY SCHOOL MARK*[1]

Colorado State College of Education

CLARENCE CARL MOORE

There are certain educational practices which have been adopted by educators and which have been questioned little by them even though these practices play a very significant part in the life of every school child The school mark is one of these The practice of awarding school marks to indicate pupil progress and educational growth is so ancient that the date of its origin has been lost Moreover, educators are aware of the fact that no other single teaching device or educational tool has been so universally adopted or applied over such a long period of time.

These marks have served many and varied purposes during their period of existence. Of recent date, the degree of their value in educational progress and the extent to which they may be used has been seriously questioned It is probable that few educational practices have been more seriously and meaningfully controverted This has led to changed views concerning the importance of the mark and the purposes to which it might be applied. The controversy concerning the mark has led to refinement in the devices by which it might be ascertained. There is still considerable controversy concerning what the mark represents.

The science of pupil measurement has had a very rapid growth during the past two decades The testing program of educational institutions has been changed, refined, and utilized in facilitating educational progress Any movement which has had as rapid growth as the objective and standardized testing program is likely to be applied in certain situations for which it is ill adapted

The oral quizz has been a part of the daily routine in classrooms from time immemorial There was a time when it was practically the whole teaching process Written examinations are more recent

---

*Received in the Editorial Office on May 20, 1938

[1] Based on Moore, C C *The Relation of Teachers' Marks to Standardized Tests* Unpublished Doctor's Field Study, Field Study No 1 Colorado State College of Education, Greeley, Colorado, 1938

than oral testing although this method of measurement has been utilized by educational procedure for many centuries. The written examination was used by Chinese educators thirteen hundred years ago. Grecian and Roman teachers were also familiar with the written examination (9). Horace Mann made an extended argument for the use of the written examination in the public schools of Massachusetts in 1845 (9). This leads to the conclusion that the oral examination had predominated previous to that date

During the early history of mental testing the idea was prevalent that if a person's native ability was accurately known, his success in school or life could be predicted with a high degree of perfection This was the natural thing to expect. Later on when a better view of the situation was obtained, it was concluded that a person of high mental ability is sometimes an indolent time waster Due to this fact a person of average native ability may excel a person of high native ability in actual accomplishment Achievement may be predicted with a certain degree of accuracy from native ability, but in a lesser degree than was thought in the beginning of the testing movement for native ability.

Rice (7) is generally given credit for having begun the standardized testing movement. The movement was an outgrowth of the desire on the part of educators to produce an instrument which would measure in the social sciences with the same exactness as the devices used for measuring in the physical sciences. When objective tests were designed, it was thought that these outcomes had been achieved. Previous to the objective testing era and contemporary with it there was much evidence produced in support of the thesis that these instruments of measurement were highly subjective and that they were, because of their subjective nature, unreliable as scientific measuring devices This brought the traditional examination into disrepute as an instrument of exact measurement and caused most of those who had previously used it to accept the objective test as an instrument more adapted to the growing needs The group testing movement developed in two directions: (*a*) The standardized commercial test and (*b*) the class-room tests constructed for teachers.

These objective tests soon convinced teachers that they were measuring with exactness. If these tests were constructed in such a manner that they had a high degree of reliability, any person using them would obtain the same results as every other person using them obtained However, a recent experiment has suggested that even though these tests are objective in the results obtained in scoring them, they are not necessarily objective in (*a*) The items chosen, (*b*) the manner in which items are stated, (*c*) the type of test item selected, (*d*) the variety and proportion of types of test items, (*e*) the grouping and other relational aspects of items, (*f*) the clarity and fullness of general and specific directions, (*g*) the personality of tester, (*h*) the time allowed for taking the test, (*i*) the number of test items, (*j*) the type of mental activity required by items, (*k*) the evaluation of pupil responses; and (*l*) pupil interpretation of items (8)

It is quite probable that many objective tests measure little except factual knowledge It may be that teachers are coming to the conclusion that desirable pupil growth consists of more than growth in the development of fact collecting skills While factual knowledge is important as a basis for forming mature judgments, it is worth little unless it is definitely related to the useful experiences of the child However, it is generally recognized that standardized-objective tests constitute an important part of the program of any educational institution, if the limits of their usefulness are generally recognized

THE PURPOSE OF THIS STUDY

It is generally agreed by a large number of educators that there are tests at the present time which measure with a high degree of accuracy the pupil's mental ability, It is almost as generally agreed that there are single and battery tests which accurately measure the achievement of pupils in certain subjects and subject fields

If it is true that tests of mental ability actually test mental ability and pupils work up to or in proportion with that ability, then, there should be a high degree of relation between the results of a test of mental ability and teacher's marks If achievement tests and teacher's marks measure the same qualities, there should be a high degree of relation between these two qualities.

The ultimate objective of this study is to determine the approximate accuracy with which teacher's marks indicate the achievement of pupils when this achievement is measured by standardized tests of achievement and the accuracy with which teacher's marks indicate the mental ability of pupils when mental ability is measured by a standardized test of mental ability

The proximate objectives approached the problem in a three-fold manner (*a*) To determine the relation between teachers' marks as given to pupils at the end of a semester and the pupil's native ability as measured by a standardized test of mental ability; (*b*) to determine the relation between teachers' marks as given pupils at the end of a semester and the pupil's achievement as measured by a standardized achievement test; and (*c*) to determine the relation between pupils' ability as measured by a standardized test of mental ability and pupils' achievement as measured by a standardized test of achievement

PROCEDURE

The field of inquiry has been limited to the data gathered in the Glenrock-Parkerton School System, Glenrock, Wyoming, the Branson Elementary Schools, Branson, Colorado; and the Grover-Hereford School System, Grover, Colorado

The fifth and sixth grades were selected for study in this problem. The New Stanford Achievement Test was selected to measure achievement and the Otis Group Intelligence Scale was selected to measure mental ability

PERTINENT STUDIES IN CORRELATION

The writer has found few studies that even approached a complete report on all the correlations of this study. Those studies which are most comparable are shown in Tables 1 and 2 These are used as a check on the accuracy of the results obtained in this study

TABLE 1

A COMPARISON OF THE RELATION BETWEEN ELEMENTARY SCHOOL INTELLIGENCE SCORES AND TEACHERS' MARKS IN ELEMENTARY SCHOOL SUBJECTS

Author	Intelligence	Teachers' marks	*N*	*r*	*PE*	*E*
Dickinson (1)	Scores	English	149	45	.04	.11
This Study	Otis Scores	Language & lit	173	67	.13	26
Dickinson (1)	Scores	Mathematics	149	29	05	04
Keites (4)	Terman *IQ*	Mathematics	112	66	04	25
This Study	Otis Scores	Mathematics	174	.59	03	19
Dickinson (1)	Scores	History	149	35	05	06
This Study	Otis Scores	Social Science	169	49	.04	.13

TABLE 2

A SUMMARY OF THE COEFFICIENTS OF CORRELATION DERIVED FROM THE CORRELATION OF THE COMPOSITE SCORES OF AN ACHIEVEMENT TEST CORRELATED WITH THE SCORES OBTAINED BY MEANS OF A PSYCHOLOGICAL EXAMINATION

Author	*N*	*r*	*PE*	*E*
Heilman (2)	838	74	01	33
Woolf (10)	157	.74	02	33
Moore*	172	69	03	28

*This study.

Dickinson (1) correlated teachers' marks in different subjects with intelligence at the eighth, ninth, and tenth grade levels. The eighth grade level corresponds most closely with the levels in this study and is shown in Table 1. The subjects which he correlated with intelligence that correspond most closely to the subjects correlated with intelligence in this study are English, mathematics, and history Kertes (4) correlated marks in eighth-grade arithmetic with intelligence This study is also shown in Table 1. Woolf (10) correlated the composite scores of an achievement test with the scores obtained by means of a psychological examination The subjects used by Woolf were 10-year old children Heilman (3) correlated the same two measures which are comparable to this study as were correlated by Woolf. However, Heilman's study was based upon over 800 cases, while Woolf's study was based on less than 200 cases The results of these studies are given in Table 2 These findings are compared with the results obtained by the writer

Table 1 shows a comparison of the relation between elementary school intelligence scores and teachers' marks in elementary school subjects The coefficients of correlation in this study are generally higher than the coefficients found by Dickinson The coefficient of correlation in mathematics and intelligence determined by Kertes is higher than the coefficient of correlation determined in this study for the same qualities

Table 2 shows a summary of the coefficients of correlation derived by correlating the composite scores on an achievement test with the scores obtained by means of a psychological examination The coefficients of correlation obtained by Woolf and Heilman are slightly higher than the coefficient of correlation determined in this study However, Table 3 shows that if the number of scores included in a

TABLE 3

SUMMARY OF THE ELEMENTARY SCHOOL COEFFICIENTS OF CORRELATION, THE SIZE OF THE GROUP IN EACH CORRELATION, THE PROBABLE ERROR OF EACH CORRELATION, THE COEFFICIENT OF DEPENDABILITY

	Measures correlated											
	Marks and achievement				Marks and intelligence				Achievement and intelligence			
Divisions	*N*	*r*	*PE*	*E*	*N*	*r*	*PE*	*E*	*N*	*r*	*PE*	*E*
Language and Lit	193	44	04	.10	173	67	05	26	172	.65	03	24
Mathematics	192	60	03	20	174	59	03	19	171	63	05	27
Reading	190	56	03	17	170	47	04	12	171	74	02	33
Social Science	173	39	04	08	169	49	04	13	172	54	04	16
Averaged Scores*	192	61	03	21	174	69	03	.28	172	69	05	28

*All four divisions of subject matter on the new Stanford Achievement Test were averaged before being correlated with the average of teachers' marks or with intelligence composite are increased and these are then correlated with some other measure, a higher coefficient of correlation will result The composite score of this study as referred to in Table 1 included only four subdivisions of the New Stanford Test while the composite scores of Woolfs study and those Heilman's study included all ten subdivisions of the same test These two studies agree rather closely with the comparable parts of this study This would suggest that the procedure and technic of this study have been definitely fitted to the purpose they propose to serve

INTERPRETATION OF TABLE 3

Column 1 of Table 3 gives the different divisions of subject matter in which the correlations were made. Three sections of the table were grouped under the term "Measures correlated" These three sections have been named "marks and achievements," "marks and intelligence," and "achievement and intelligence" The term "marks" here applies to those measures of achievement as are given to the pupils at the end of a semester by the teachers The term "achievement" is here applied to the results as obtained by a standardized achievement test. The term "intelligence" is here applied to those scores which pupils make when a standardized test of mental ability is administered to them.

Columns 2, 6, and 10 give the number of cases in each correlation Columns 3, 7, and 11 give the size of the coefficient obtained from each correlation Columns 4, 8, and 12 show the probable error of each coefficient Columns 5, 9 and 13 show the coefficient of dependability for each correlation

The coefficients of correlation are much more uniform in size for marks and achievement in this table than were the corresponding coefficients for high school groups (6)

TEACHERS' MARKS AND ACHIEVEMENT

The largest coefficient of correlation is for the average test scores and teachers' marks Of the four subjects correlated with achievement, mathematics is first, reading is second, language and literature is third, and social science is fourth The lowest correlation of this group is almost 10 times its probable error. This suggests a true relationship for each of these coefficients

When these data are classified according to the coefficient of dependability, the following conditions are found to exist: The correlation for teachers' marks in language and literature and achievement test scores in language and literature is 10 per cent of the way between a best guess (2) and perfect correlation, the correlation for teachers' marks in mathematics and achievement test scores in mathematics is 20 per cent of the way between a best guess and perfect correlation; the correlation for teachers' marks in reading and achievement test scores in reading is 17 per cent of the way between a best guess and perfect correlation, the correlation for teachers' marks in social science and achievement test scores in social science is 8 per cent of the way between a best guess and perfect correlation; and the correlation for teachers' marks in four subjects and the average achievement scores in four divisions of subject matter is 21 per cent of the way between a best guess and perfect correlation

TEACHERS' MARKS AND INTELLIGENCE

The coefficients of correlation for teachers' marks and intelligence test scores are in general slightly higher than the coefficients of correlation for marks and achievement.

When the data of the second section of Table 3 are classified according to the coefficient of dependability, the following conditions are found to exist: The agreement for teachers' marks in language and literature and intelligence test scores is 26 per cent of the way between a best guess and perfect correlation, the agreement for teachers' marks in mathematics and intelligence test scores is 19 per cent of the way between a best guess and perfect correlation; the agreement for teachers' marks in reading and intelligence test scores is 12 per cent of the way between a best guess and perfect correlation; the agreement for teachers' marks in social science and intelligence test scores is 13 per cent of the way between a best guess and perfect correlation, and the agreement for teachers' marks in the average for four subjects and intelligence test scores is 28 per cent of the way between a best guess and perfect correlation.

ACHIEVEMENT AND INTELLIGENCE

The correlations for achievement test scores as measured by a standardized test and for intelligence test scores as measured by a standardized test are generally higher than the correlations presented in the other two sections of this table. In this section is found the highest relationship of the entire study There is a coefficient of correlation of 74 between achievement in reading and intelligence test scores. The second and third highest correlations exclusive of those found by correlating averaged scores with some other quality is also found in this section.

The coefficient of dependability shows the following conditions to exist· The agreement for achievement test scores in language and literature is 24 per cent of the way between a best guess and perfect correlation, the agreement for achievement test scores in mathematics and intelligence test scores is 27 per cent of the way between a best guess and perfect correlation; the agreement for achievement test scores in reading and intelligence is 33 per cent of the way between a best guess and perfect correlation, the agreement for achievement test scores in social science and intelligence is 16 per cent of the way between a best guess and perfect correlation; and the agreement for the average of achievement test scores in the four subjects and intelligence is 28 per cent of the way between a best guess and perfect correlation

Conclusions

1 These data definitely show that there is not a high degree of relationship between the marks that teachers assign their pupils at the end of a semester and the standing of pupils on either standardized tests of achievement or intelligence.

2. Care should be exercised to discriminate between the term "marks," when applied to express the teacher's estimate of the pupil's achievement, and the term "achievement," when applied to express the pupil's achievement as determined by standardized tests. These two measures differ considerably

3 The study suggests that more attention should be given to devices which are used by teachers in measuring pupils. A preliminary study of the true-false test points out the extent of some of the errors produced by this device (5).

4. The study also suggests that educators should give considerable attention to the situations in which standardized test results may give adequate information Some are limited to a general survey of pupil achievement while others sample more definitely

specific interests. However, most standardized tests of achievement do not sample adequately the fields of the pupils specific interests to give an adequate picture of the whole child

REFERENCES

1 DICKINSON, C E Reading ability and school achievement. *School Rev*, 1925, **33**, 616-625

2 HEILMAN, J D The *k* and *g* methods of interpreting the coefficient of correlation *J. Educ Psychol.*, 1937, **28**, 232-236

3 ——— Relative influence upon educational achievement of some hereditary and environmental factors *Twenty-seventh Year-book, Nat. Soc. Stud Educ*, 1928

4 KERTES, F Ability grouping in high school *Math Teach*, 1932, **25**, 5-16

5. MOORE, C C Factors of chance in the true-false examination *J Genet Psychol.*, 1938, **53**, 215-229

6 ——— Needed' A new plan of grading. *School Exec*, 1937, **56**, 402-403

7. ODELL, C W *Educational Measurement in High School* New York Century, 1930

8. PULLIAS, E V. Variability in Results from New-Type Achievement Tests Durham, N C Duke Univ Press, 1937

9. RUCH, G M The Objective or New Type Examination Chicago, Scott, Foresman, 1929

10. WOOLF, HENRIETTA K The relative influence of some environmental factors and mental age or educational age (Unpublished master's thesis, Colorado State College of Education)

Gill Public Schools
Gill, Colorado

The Journal of Genetic Psychology, 1939, 54, 295-312

A COMPARATIVE STUDY OF THE DIFFICULTY OF THE 100 ADDITION COMBINATIONS*

State Teachers College, Johnson City, Tennessee

LESTER R. WHEELER

An evaluation of the methods used in studying the relative difficulty of learning the addition combinations depends upon the degree to which investigators have succeeded in isolating inherent difficulty from the variable conditions of the learning process Eight different methods have been used in determining the relative difficulty of the combinations.

1 Questionnaire, teachers' judgment, and other introspection methods,
2 Comparison of results obtained from normal and backward children,
3 Comparison of the number of errors for each combination, the most difficult being the one where the largest number of errors occur
4 Reaction time, measuring the time required to give the sum of each combination
5 Measuring the retention after a period of non-use
6 Degree of difficulty in learning as indicated by trials or repetitions necessary for accurate and immediate recall
7 Familiarity with the combinations among children at the time of school entrance
8 Composite difficulty rankings based upon combining various investigations

The effect of unequal practice on learning has been neglected in all investigations except the two which combined several of the above methods Holloway's (14) published in 1915, and that of Knight and Behrens (17), 1928, commonly known as the *K-B* study The latter outstanding contribution furnishes adequate grounds for an attack on Clapp's more extensive investigation and others which overlooked the effect of practice Knight and Behrens recorded the progress of each child while learning the hundred

---

*Received in the Editorial Office on June 3, 1938

addition facts, and analyzed the data on the basis of learning, reaction-time, number of errors, and retention Their results disagree with previous studies in several respects and emphasize the importance of equalizing trials, exposures, repetitions, and practice. The long and laborious technique used by Knight and Behrens, while sound, is limited to the study of a small number of cases. With the new type of teaching materials, *ADD-O, COUNT-O,* and *MULT-O,*[1] it is possible to study the relative difficulty of the fundamental number combinations during the actual learning period with a large number of cases This article presents the results of the study of the relative difficulty of the hundred addition facts when taught by the game of *ADD-O*, and compares the final ranking with previous investigations

ADD-O is an arithmetic game that teaches and drills the hundred addition combinations through a free-play activity. Frequency of repetition of the combinations, vividness of presentation, typographical technicalities, length of learning time, type and quality of motivation, degree of satisfaction, variations within the class due to individual differences, and method of teaching are controlled by the nature of the *ADD-O* material. The correct combinations and answers are always before the children, minimizing the possibilities of establishing incorrect responses. Accuracy is checked and speed is regulated in an effort to control undesirable habits of computation. The game is simple and, even though motivated by a real spirit of play, attention is focused on the number combinations rather than on the process of the game Interest in the game increases with use, the more combinations a child learns the more fun it is to play[2]

Data for this experiment were gathered during 20 consecutive school days in October and November, 1936. Tests were given to 125 children in Grade II*B* in four public schools in Johnson City, Tennessee, and in the Training School of the East Tennessee State Teachers College The median *IQ* for the total group was 103 ± 996 A mimeographed test of the combinations in both equation and column form was given to determine which combinations were familiar to the children. During the following 10 days *ADD-O was played 20 minutes a day*, then the same test *repeated.*

---

[1] Augsburg Publishing Co., Morristown, Tennessee, 1936-1937

[2] For details of game, see Goforth (11).

At the end of 20 days the final test showed that 25 per cent of the children had mastered over 80 of the combinations. No incentive other than the game itself was used for motivation, and no attempt was made to supervise or assist the teachers in any way other than to administer the tests and instruct them how to play the game The repetition of the combinations and various other factors of the learning process were controlled during the arithmetic period by the objective nature of the game, but not checked as to incidental uses in the other school-room activities. Concepts of the adding process had been taught before this experiment was begun.

Table 1 shows the results of Tests I, II and III in both the equation and column forms of the hundred combinations In this table those who approached within ten points of 100 per cent mastery on the second test and maintained their high score on the third test were eliminated in order to get a true picture of the rate of learning, thus reducing the number of cases in Test III to 111. The initial test shows that these second grade children knew on an average of about 10 or 12 combinations at the beginning of the experiment

TABLE 1

RESULTS OF TESTS I, II, AND III IN BOTH EQUATION AND COLUMN FORMS

	No of cases	Median	*Q*-1	*Q*-3	*Q*
Test I					
Equation Form	125	12 95±1 91	1 19	35 47	17 14
Column Form	125	10 68±1.56	1 88	29.81	13 97
Difference		2 27±2 46	— 69	5 66	3 17
Diff P E (diff)		92			
Chances in 100		73			
Test II					
Equation Form	125	45 35±2 50	22 25	67 29	22 52
Column Form	125	35 83±2 11	18 91	56 75	18 92
Difference		9 52±3 27	3 34	10 54	3 60
Diff P E (diff)		2 91			
Chances in 100		97			
Test III					
Equation Form	111	61 25±2 79	37 34	84 37	23 52
Column Form	111	55 41±2 77	34 75	81 56	23 41
Difference		5 84±3 93	2 59	2 81	11
Diff P E (diff)		1 23			
Chances in 100		79			

At the end of 10 days they knew 45 when presented in the equation form and 35 in the column form At the end of 20 days their achievement reached 61 and 55 respectively. Similar progress is shown by the first and third quartiles.

The *PE* (*diff.*) and chances in 100 indicate a significant difference in favor of the column form for Test II, but a smaller difference for Tests I and III The increase in the difference between the equation and column forms for Test II is probably explained by the fact that all the flashcards in the *ADD-O* material are printed in the equation form although both forms are equally distributed on the child's playing cards While this difference is not large, the trend appears significant enough during the first 10 days to suggest that the form of presenting the combinations may have some influence on learning, although during the second 10 days there is an indication that more transfer has taken place.

The rate of learning is given in Table 2 As the experiment progressed there was a slight decrease in the rate of mastering the combinations, probably due to the increasing difficulty rather than to the elimination of the pupils who mastered the combinations on Test II The correlation between *IQ* and rate of learning is very low, $r = .23 \pm 033$ During the 20 days of the experiment the children gained on an average of nearly two and a half combinations a day

TABLE 2
RATE OF GAIN IN LEARNING THE 100 ADDITION COMBINATIONS

Number of combinations learned	During ten days	During twenty days
1		
Equation Form	32 40	48 30
Column Form	25 15	44 73
Average	28 78	46 51
2 Average Number of combinations learned per day		
Equation Form	3 24	2 42
Column Form	2 52	2 24
Average	2 83	2 38

The combinations were ranked from easiest to most difficult on the basis of the percentage of children mastering each on each of the three tests Correlations between the difficulty rankings on the three tests are shown in Table 3 While there is a high correlation between the difficulty ranks of all three tests, there is also a definite trend indicating that as the learning of the combinations progressed the relationship of the difficulty rankings became closer with a very high correlation between Tests II and III The general trend in comparing the equation and column forms shows practically no difference on the final test

TABLE 3

INTERCORRELATIONS OF THE DIFFICULTY RANKING OF THE THREE TESTS, AND A COMPARISON OF THE DIFFICULTY OF THE EQUATION AND COLUMN FORMS

1	Equation Form			
	Test I with Test II	r *=	84±	021
	Test I with Test III		73±	033
	Test II with Test III		95±	007
2	Column Form			
	Test I with Test II		89±	014
	Test I with Test III		80±	025
	Test II with Test III		92±	011
3	Equation Test I with Column Test I		54±	057
	Equation Test II with Column Test II		84±	021
	Equation Test III with Column Test III		94±	008

*rank difference method

The rate of learning and the correlations of the three difficulty ranks indicate that the results of Test III, equation form, probably give us the best rank of the relative difficulty of the hundred combinations, this final ranking is shown in Table 4 The easiest combination is 0+0, mastered by 96 per cent of the group, and the difficulty increases gradually to 7+9 which was learned by only 23 per cent A study of the increment of difficulty existing between the combinations shows the difference from one rank to the next is consistently small throughout the table; no great amount of difficulty exists between rank one and two, two and three, etc.

This difficulty ranking is compared with those of other investigators in Table 5. MacLatchy (20) studied the degree of familiarity which children have with 20 selected combinations at the time of entering school. As her method of study appears to allow computation by "counting," one suspects that her results measure familiarity with the concept of addition rather than mastery of the combination

TABLE 4

THE 100 ADDITION COMBINATIONS RANKED FROM EASIEST TO MOST DIFFICULT

	Rank	% of cases mastering each	Step difference in rank		Rank	% of cases mastering each	Step difference in rank
0+0	1.0	96.0	—	5+3	50.5	61.6	0.0
0+1	2.5	90.4	5.6	5+2	52.0	60.8	0.8
0+5	2.5	90.4	0.0	5+4	53.0	59.2	1.6
0+8	4.0	86.4	4.0	3+5	54.0	58.4	0.8
4+4	6.0	84.8	1.6	9+9	55.5	57.6	0.8
9+0	6.0	84.8	0.0	4+2	55.5	57.6	0.0
5+0	6.0	84.8	0.0	7+2	59.5	56.8	0.8
1+0	9.0	84.0	0.8	4+3	59.5	56.8	0.0
1+1	9.0	84.0	0.0	2+9	59.5	56.8	0.0
2+0	9.0	84.0	0.0	6+6	59.5	56.8	0.0
6+0	13.0	83.2	0.8	3+7	59.5	56.8	0.0
1+2	13.0	83.2	0.0	3+9	59.5	56.8	0.0
0+4	13.0	83.2	0.0	2+8	63.0	56.0	0.8
3+3	13.0	83.2	0.0	9+2	64.0	55.2	0.8
3+0	13.0	83.2	0.0	7+3	65.0	54.4	0.8
8+0	17.0	82.8	0.4	3+8	66.0	53.6	0.8
0+7	17.0	82.8	0.0	2+6	67.5	52.0	1.6
0+6	17.0	82.8	0.0	3+6	67.5	52.0	0.0
1+9	19.5	81.6	1.2	6+4	69.5	51.2	0.8
1+3	19.5	81.6	0.0	9+3	69.5	51.2	0.0
6+1	22.0	80.8	0.8	6+3	71.0	49.6	1.6
7+0	22.0	80.8	0.0	4+5	72.0	48.8	0.8
4+0	22.0	80.8	0.0	5+6	73.0	47.2	1.6
0+2	24.5	80.0	0.8	7+4	74.5	46.4	0.8
4+1	24.5	80.0	0.0	8+8	74.5	46.4	0.0
1+8	26.5	78.4	1.6	8+4	76.0	45.6	0.8
2+1	26.5	78.4	0.0	4+7	78.0	43.2	2.4
3+1	28.0	77.6	0.8	6+5	78.0	43.2	0.0
1+4	29.5	76.8	0.8	4+9	78.0	43.2	0.0
0+9	29.5	76.8	0.0	9+4	80.0	42.4	0.8
1+5	31.0	76.0	0.8	4+8	81.0	40.8	1.6
0+3	32.0	75.2	0.8	5+9	82.0	40.0	0.8
1+7	33.0	74.4	0.8	6+9	83.0	35.1	4.9
2+3	34.0	73.6	0.8	8+5	84.0	33.6	1.5
5+1	35.5	72.8	0.8	7+6	85.5	32.8	0.8
8+1	35.5	72.8	0.0	9+5	85.5	32.8	0.0
5+5	37.5	72.0	0.8	5+7	87.0	32.0	0.8
2+2	37.5	72.0	0.0	9+6	89.0	31.2	0.8
9+1	39.0	70.4	1.6	7+5	89.0	31.2	0.0
3+4	40.5	69.6	0.8	4+6	89.0	31.2	0.0
3+2	40.5	69.6	0.0	7+8	91.0	30.4	0.8
2+5	43.0	68.8	0.8	8+6	92.0	28.8	1.6
1+6	43.0	68.8	0.0	9+8	94.0	28.0	0.8
7+1	43.0	68.8	0.0	6+7	94.0	28.0	0.0
2+7	45.0	65.5	3.3	9+7	94.0	28.0	0.0
8+3	46.5	64.8	0.7	6+8	96.0	26.4	1.6
7+7	46.5	64.8	0.0	5+8	97.5	24.8	1.6
8+2	48.0	64.0	0.8	8+7	97.5	24.8	0.0
6+2	49.0	62.4	1.6	8+9	99.0	24.0	0.8
2+4	50.5	61.6	0.8	7+9	100.0	23.2	0.8
					Average		0.73

TABLE 5

COMPARISONS WITH OTHER INVESTIGATIONS

	No of Combinations	
MacLatchy with Wheeler (Test I)	20	r= 68± 084
Phelps with Wheeler (Test III)	54	67± 052
Clapp with Wheeler (Test III)	100	69± 037
K-B with Clapp (*K-B* data)	100	*66±.040*
K-B with Wheeler (Test III)	100	92± 011

This might influence her conclusion that the easiest combinations are those in which "1" is added to a larger number, followed in order of difficulty by "2," "3," and "4" Phelps (23) analyzed data based on the number-of-errors made by 270 eighth grade pupils in the 55 combinations, and Clapp (8) used the same method with 7000 pupils in Grades III through VIII. It is interesting to note the very high correlation between our ranking and that of the *K B* investigation, but less agreement with the others which were based on methods that disregarded the variable factors in the learning process

Further comparisons may be made by studying the overlapping of the various rankings to discover where the points of greatest difference occur in these lists The combinations in the Wheeler list having a greater rank difficulty of 20 or more places than is found in the *K-B* list are:

0+3, 31 points difference in rank
2+2, 27 5 " " " "
5+1, 24 5 " " " "
7+1, 31 " " " "

The combinations in the *K-B* list having a greater rank difficulty of 20 or more places than is found in the Wheeler list are.

0+8, 26 points difference in rank
4+4, 35 " " " "
4+0, 25 " " " "
1+2, 20 " " " "
5+3, 19 5 " " " "
3+4, 25 5 " " " "
3+5, 27 " " " "

Twenty combinations in the Wheeler list have a rank difficulty of 20 to 67 points greater than is found in the Clapp list, and 18 combinations in the Clapp list are 20 to 68 points greater than in the Wheeler list Out of 20 combinations in Clapp's list having a greater difficulty of 20 or more points than in the *K-B* ranking, 12 of these are the same as appear in the Clapp Wheeler comparison: 2+3, 8+0, 1+2, 1+9, 0+5, 0+6, 6+0, 0+4, 0+1, 9+0, 1+0, 0+3. Notice that with three exceptions all these are the zero combinations, which were probably slighted in the teaching program for Clapp's group of children.

The frequency of occurrence of the combinations in the *ADD-O* games correlated with our difficulty rank shows that there is no relationship between the number of times the combination is repeated and the ease with which it is mastered, $r = -09 \pm .013$ This indicates that repetition alone is not sufficient to insure mastery of the hundred combinations, and agrees with other psychological experiments attacking the validity of the laws of learning.

According to Clapp there is a negative correlation between the difficulty of a combination and that of its reverse, as shown in Table 6. The data of the *K-B* and this investigation agree that there is a positive relationship between the difficulty of a combination and its reverse, although the correlations indicate only the general trend, neither correlation is high enough to predict that such a relationship would hold for every combination The average displacement, or difference in rank points between a combination and its reverse, is practically the same for both the *K-B* and the Wheeler data, but nearly twice as much in Clapp's study Again, there appears a close agreement between the *K-B* study and our investigation.

TABLE 6

COMPARISON OF THE RANK ORDER OF DIFFICULTY OF THE 45 COMBINATIONS WITH THEIR REVERSES

1 Relation of 45 combinations to their reverses	
Clapp	$r = -38 \pm .013$
K-B	$61 \pm .009$
Wheeler	63 ± 009
2 Average displacement or average difference in rank between direct and reverse combinations	
Clapp	17 33
K-B	9 04
Wheeler	9 64

Table 7 shows that Clapp found a greater dispersion in the diffi-

TABLE 7

COMPARISONS OF RANK ORDER OF DIFFICULTY IN *K-B*, CLAPP AND WHEELER INVESTIGATIONS

	Average rank order			Average rank difference	
	K-B	Clapp	Wheeler	*K-B* & Wheeler	Clapp & Wheeler
1. Zero combinations					
Direct	17.2	38.3	11.8	10.4	27.3
Reverses	14.8	49.4	16.0	10.6	30.5
2. Doubles combinations	31.7	11.6	34.0	13.0	25.6

culty ranking of the zero combinations than did either *K-B* or Wheeler. Thirteen of the zero combinations in Clapp's study fall between ranks 31 and 60. *K-B* reports 17 of the 19 zero combinations of a rank difficulty of 30 or less, and our data show 18 of the combinations ranking under 30. The range of difficulty for the zero combinations was, according to Clapp, 1 to 75, according to *K-B, 1 to 47, and according to Wheeler 1 to 32.* The average rank difference between the Wheeler and Clapp data is three times greater than between Wheeler and *K-B*. In both the *K-B* and Wheeler studies the zero combinations were given the same teaching emphasis as the other combinations, and the data indicate that the method of teaching definitely influences the difficulty ranking of the zero combinations.

Is the relative difficulty of the combinations explained by the size of the sums? Knight and Behrens believe it is, and in Table 8 we have taken some of their data for purposes of comparison. The *K-B* Size of Sum hypothesis (Column 6) may be found by ranking the sums after the combinations have been arranged according to the size of their sums. It is interesting to note that our difficulty rank (Column 3) is closer to the *K-B* hypothesis than the *K-B* ranking (Column 7).

In order to find out whether any one group of combinations has a tendency to throw the difficulty ranking out of line, we compared our ranking with the *K-B* hypothesis when the doubles were eliminated (Column 4) and when both the doubles and zero combinations were eliminated (Column 5). A summary of the results is given in Table 9. It appears that eliminating the zero combinations does not bring our ranking any closer to the size-of-sum hypothesis,

TABLE 8

SIZE-OF-SUM HYPOTHESIS AND ITS RELATION TO VARIOUS DIFFICULTY RANKS

(1) Sum	(2) Frequency rank of sums	(3) Av of rank difficulties	(4) Av rank of difficulties eliminating doubles	(5) Av rank of difficulties eliminating doubles and zeros	(6) Approximate rank difference expected by the *K-B* Size-of-Sum Hypothesis	(7) Av rank of difficulties as reported by *K-B* investigation
0	3.0	1.0	—	—	1.0	17.0
1	1.5	2.5	2.5	—	2.5	4.5
2	5.0	14.0	16.5	—	5.0	9.3
3	6.0	21.0	21.0	19.5	8.5	17.0
4	10.0	24.0	20.5	23.5	13.0	27.2
5	9.0	22.8	22.8	32.0	18.5	31.2
6	14.5	30.7	33.7	43.0	25.0	24.7
7	17.5	37.4	37.4	43.3	32.5	38.1
8	14.5	36.2	38.5	49.5	41.0	47.0
9	19.0	46.5	46.5	53.8	50.5	43.5
10	14.5	54.4	56.5	56.5	60.0	54.6
11	17.5	67.4	67.4	67.4	68.5	60.0
12	14.5	74.1	76.5	76.5	76.0	72.0
13	12.0	86.5	86.5	86.5	82.5	89.2
14	11.0	80.4	89.0	89.0	88.0	82.0
15	7.5	90.0	90.0	90.0	92.5	87.3
16	7.5	89.1	97.0	97.0	96.0	84.0
17	4.0	96.5	96.5	96.5	98.5	92.5
18	1.5	55.5	—	—	100.0	62.0

TABLE 9

AVERAGE DIFFERENCE BETWEEN DIFFICULTY RANKINGS AND THE *K-B* SIZE-OF-SUM HYPOTHESIS

		Average number of points difference
1	*K-B* rank, all combinations included	6.60
2	Wheeler rank, all combinations included	4 09
3	Wheeler rank, eliminating doubles and zeros	5 50
4.	Wheeler rank, eliminating doubles	3 94

but there is a slight advantage in eliminating the doubles. It seems possible, therefore, that the doubles combinations are the ones which draw the arrangment out of line. Notice in Table 8 that 0+0 and 9+9 do not conform to hypothetical arrangement in the *K-B* difficulty ranking, and the latter is out of place in the Wheeler list

The greatest differences between the *K-B* hypothesis and our difficulty ranking appear in the combinations with sums of 2 to 8 Eliminating the zero combinations does not help this any, but taking out the doubles reduces this difference from 70 to 40 points. Knight and Behrens suggest that the following factors may cause discrepancies between the difficulty ranks and hypothesis

1 Hypothesis may be wrong.
2. Data may be imperfect.
3 Positions in the learning series may affect the arrangement, as the smaller combinations are usually taught first, the children may have profited in later combinations by general experience.
4 Frequency of repetition in drill

Our data show that the last two factors are not determining elements We have shown that repetition alone does not affect the relative difficulty, and the technique used in formulating the difficulty ranking eliminates the possibility of any appreciable transfer taking place from general experience A study of the rate of learning indicates that the easiest combinations are learned first and more rapidly than the more difficult ones If general experience is a factor, then the difficult combinations are actually even more difficult than the relative ranking indicates.

While there is a definite relationship between the arrangement of the combinations according to size-of-sums and the relative difficulty ranking, $r = .85 \pm .019$, the size of the sum does not determine the relative difficulty as Knight and Behrens have suggested. The correlation of the frequency of occurrence of the sums in the teaching materials with the average of the difficulty rankings for each sum shows that no relationship exists, $r = .06 \pm .103$ (see Columns 2 and 3, Table 8). The *K-B* investigation mentions the possibility that the difficulties may arrange themselves in groups when one of the two addends is held constant The authors conclude:

> There is some evidence for suspecting an arrangement of difficulty of the 100 combinations organized in such a way to show considerable influence of this hypothesis We are inclined to think, however, that the Size-of-Sum Hypothesis in and of itself takes care of all explanation which Hypothesis 2 (grouping according to constant addend) is capable plus more of the total explanation on its own account (17, p 59)

In Table 10, the grouping of the combinations according to the size of the common addends indicates that the suspicions of Knight and Behrens were well founded, and the size of the sum appears to be the result rather than the cause of the difficulties The constant addend arrangement does of necessity overlap the size-of-sum hypothesis; if the difficulties are ranked according to one arrangement, they will also fall into the arrangement of the other, and vice versa

By comparing the averages of the groups, we find that the combinations which contain a common addend, in general, appear to be of similar difficulty, and the difficulty increases with the increase in the size of the common addend. A study of the overlapping of the various groups help to substantiate this general trend. Table 11 gives the number of combinations in one group which exceed the average difficulty of the group immediately following, and the amount of overlapping We find here that only four out of the 100 combinations are out of line according to this grouping They are the combinations 0+3, 6+3, 4+6, and 5+8, and the average number of points overlapping is small.

Does it make any difference in degree of difficulty whether the second addend is larger or smaller than the first? From Table 10 we find that when the second addend is smaller than the first the average difficulty rank is 66.7, and when the second addend is larger

TABLE 10

THE 100 ADDITION COMBINATIONS GROUPED ACCORDING TO THE SIZE OF THE ADDENDS AND REVERSES

	Combination	Difficulty rank	Reverse	Difficulty rank	Size of sum
(0)	1+0	9.0	0+1	2.5	1
	2+0	9.0	0+2	24.5	2
	3+0	13.0	0+3	32.0	3
	4+0	22.0	0+4	13.0	4
	5+0	6.0	0+5	2.5	5
	6+0	13.0	0+6	17.0	6
	7+0	22.0	0+7	17.0	7
	8+0	17.0	0+8	4.0	8
	9+0	6.0	0+9	29.5	9
	Average	13.0		16.0	5
	Range	6 - 22		2.5 - 29.5	
(1)	2+1	26.5	1+2	13.0	3
	3+1	28.0	1+3	19.5	4
	4+1	24.5	1+4	29.5	5
	5+1	35.5	1+5	31.0	6
	6+1	22.0	1+6	43.0	7
	7+1	43.0	1+7	33.0	8
	8+1	35.5	1+8	26.5	9
	9+1	39.5	1+9	19.5	10
	Average	31.8		24.4	6.5
	Range	22 - 39.5		19.5 - 43	
(2)	3+2	40.5	2+3	34.0	5
	4+2	55.5	2+4	50.5	6
	5+2	52.0	2+5	43.0	7
	6+2	49.0	2+6	67.5	8
	7+2	59.5	2+7	45.0	9
	8+2	48.0	2+8	63.0	10
	9+2	64.0	2+9	24.0	11
	Average	52.6		46.7	8.0
	Range	40.5 - 64		24 - 67.5	
(3)	4+3	59.5	3+4	40.5	7
	5+3	50.5	3+5	54.0	8
	6+3	71.0	3+6	67.5	9
	7+3	65.0	3+7	59.5	10
	8+3	46.5	3+8	66.0	11
	9+3	69.5	3+9	59.5	12
	Average	60.3		57.8	9.5
	Range	46.5 - 71		40.5 - 67.5	
(4)	5+4	53.0	4+5	72.0	9
	6+4	69.5	4+6	89.0	10
	7+4	74.5	4+7	78.0	11
	8+4	76.0	4+8	81.0	12
	9+4	80.0	4+9	78.0	13

TABLE 10 — (*continued*)

	Combination	Difficulty rank	Reverse	Difficulty rank	Size of sum
	Average	70.6		79.6	11.0
	Range	53 - 80		72 - 89	
(5)	6+5	78.0	5+6	73.0	11
	7+5	89.0	5+7	87.0	12
	8+5	84.0	5+8	97.5	13
	9+5	85.5	5+9	82.0	14
	Average	84.1		84.9	12.5
	Range	78 - 89		73 - 97.5	
(6)	7+6	85.5	6+7	94.0	13
	8+6	92.0	6+8	96.0	14
	9+6	89.0	6+9	83.0	15
	Average	88.8		91.0	14.0
	Range	85.5 - 92		83 - 96	
(7)	8+7	97.5	7+8	91.0	15
	9+7	94.0	7+9	100.0	16
	Average	95.8		95.5	15.5
	Range	94 - 97.5		91 - 100	
(8)	9+8	94.0	8+9	99.0	17
		Doubles	0+0	1.0	0
			1+1	9.0	2
			2+2	37.5	4
			3+3	13.0	6
			4+4	6.0	8
			5+5	37.5	10
			6+6	59.5	12
			7+7	46.5	14
			8+8	74.5	16
			9+9	55.5	18
			Average	34.0	9.0
			Range	1 - 74.5	

than the first the average difficulty is 66.1. This is in agreement with the findings of Knight and Behrens that there is probably little difference whether the larger number is added to the smaller or the smaller to the larger and helps to substantiate the trend found in Table 7.

The doubles difficulties fall into a category of their own, ranging in difficulty from Rank 1 through 74.5. The average of their difficulties place them as a group immediately following the 1's, yet

TABLE 11

OVERLAPPING OF DIFFICULTIES WHEN COMBINATIONS ARE GROUPED ACCORDING TO COMMON ADDENDS

Grouping according to common addend	Number of direct combinations exceeding av of next group	Number of points exceeding	Number of reverse combinations exceeding av of next group	Number of points exceeding
"0"	0	0.00	1	5.10
"1"	0	0.00	0	0.00
"2"	0	0.00	0	0.00
"3"	1	6.70	0	0.00
"4"	0	0.00	1	6.60
"5"	0	0.00	1	1.90
"6"	0	0.00	0	0.00
Average		0.96		1.94

there are three combinations exceeding the average of the *2*'s by 7.06 points. As only one of these doubles combinations exceeds the average of the *3*'s and *4*'s, we can say that they probably belong as a group among the easier half of the combinations.

By ranking the hundred combinations according to the size of the addends, we have found the approximate difficulty of each of the groups which we might expect from the size-of-addend hypothesis. Table 12 compares the actual difficulty ranks with this hypothetical

TABLE 12

THE RELATIONS OF DIFFICULTY RANKINGS TO THE SIZE-OF-ADDEND HYPOTHESIS

Addend	Approximate difficulty rank expected by *Size-of-Addend* Hypothesis	Average of rank difficulties	Average of rank difficulties eliminating doubles
0	10.0	13.7	14.4
1	22.1	28.1	28.1
2	44.0	48.7	49.7
3	58.0	55.5	59.1
4	70.0	68.8	75.1
5	80.0	79.3	84.5
6	88.0	85.6	89.9
7	94.0	85.8	95.7
8	98.0	89.2	96.5
9	100.0	55.5	——
Average difference between hypothesis and ranking		8.3	3.5

arrangement. Again, by eliminating the doubles, there is a close agreement between our difficulty rank and the hypothesis, $r = 95 \pm .014$ The difficulty rank is somewhat closer to the size-of-addend than to the size-of-sum hypothesis, the correlation of size-of-sum with difficulty, eliminating the doubles, is 87 ± 018, and retaining the doubles $r = 85 \pm .019$ The doubles difficulties correlate 85 ± 061 with both hypotheses

Several educational implications are suggested by the results of this investigation. The combinations may be taught successfully by informal methods such as those used in the game of *ADD-O* Other studies based on similar techniques have shown that children learn more and retain longer facts indirectly presented through a free-play activity (30, 31) Best results are obtained when the mechanics of instruction are hidden from the child and when the various factors influencing learning are so integrated into an activity that learning takes place incidentally as the means to the end rather than as an end in itself This principle of indirect teaching is the same as that so successfully used in the psychology of advertising No one of the so-called "Laws of Learning" appears adequate to insure mastery Our results indicate that isolated repetition has little influence on the ease with which the combinations are learned, but repetition is probably very important when integrated with other factors It seems highly probable that learning depends upon a functional pattern, or Gestalt, made up of many interesting elements rather than upon the dominance of any single factor.

As the form in which the combinations are presented appear to influence slightly the difficulty of learning, we may conclude that the child should have the opportunity to learn both the equation and column forms, and he should be taught the reverses along with the direct combinations The teacher should not assume that the zero combinations will be learned by inference, they should be taught as any of the other hundred combinations If we assume that a child should learn the cominations according to difficulty, the size of the addend, in general, is a better index to difficulty than is the size of the sum The arrangement of the combinations in Table 10 suggests a convenient division for teaching on this basis, with doubles falling respectively at the beginning of each group according to the size of the addends The teacher should expect a decrease in the rate of learning with an increase in the difficulty of the combinations and should provide a longer learning period for the more difficult combinations.

The difficulty ranking in this and other investigations leads us to suspect that possibly the children are not memorizing the addition facts automatically so that they can give an immediate response to the sum of a combination As the size of the addend seems to be the general factor in causing differences in the difficulty ranking, we wonder if the children are not computing the sums by physical or mental counting, a crutch which is probably developed in the child while building the number concepts Psychologically the child should be able to learn $5+4=9$ as easily as $2+3=5$, but this is not the case according to the investigations of combination difficulties It might be more economical first to teach the child to memorize the combinations, and later develop the number concepts

REFERENCES

1 ANSPAUGH, G E, & PHIPPS, G C Standards of Achievement for Sixth A Grade in Common and Decimal Fractions and in Number Combinations together with an Analysis of Errors Made in These Operations Third Yearbook, Chicago, Principal's Club, 1928 Pp 54-93.

2 BATSON, W H & COMBELLICK, O E Relative difficulty of number combinations in addition and multiplication *J Educ Psychol*, 1925, **16**, 467-481

3 BEITO, E A. & BRUECKNER, L J A Measurement of Transfer in the Learning of Number Combinations Report of the Society's Committee on Arithmetic *29th Yearbook, Nat Soc Stud Educ* Bloomington, Ill Pub Sch Pub Co, 1930 Pp 569-587

4 BROWNE, C E The psychology of the simple arithmetic processes. A study of certain habits of attention and association *Amer J Psychol*, 1906, **17**, 1-37

5 BUCKINGHAM, B R Arithmetic at the crossroads *Educ Res Bull*, 1927, **6**, 2-5

6 ——— The difficulty of the number combinations *J Educ Res*, 1925, **12**, 387-391

7 BUSWELL, G T, & BUSWELL, J L Diagnostic studies in arithmetic *Suppl Educ Monog*, 1926, No 30. Pp 212

8 CLAPP, F L The number combinations Their relative difficulty and the frequency of their appearance in textbooks *Bur Educ Res*, Bull Nos 1 and 2, Univ Wisconsin, 1924 Pp 20-126

9 COUNTS, G S Arithmetic tests and studies in the psychology of arithmetic *Suppl Educ Monog*, No 4, Univ Chicago, 1917

10 FINNERAN, E M Text-book analysis-addition *Amer Educ*, 1927, **30**, 315-17

11 GOFORTH, LILLIAN A classroom experiment in teaching reading and arithmetic through games *Educ Meth*, 1938, **17**, 231-35

12. Griggs, A O The pedagogy of mathematics. *Ped Sem.,* 1912, **19**, 360

13. Heilman, J. D & Shultis, F W A study in addition *Res Bull.,* No 1, Colorado State Teachers College, Greeley, 1916 Pp 16

14. Holloway, H V An Experimental Study to Determine the Relative Difficulty of the Elementary Number Combinations in Addition and Multiplication Doctor's Thesis, Univ. Penna, 1914. Pp 102. State Gazette Pub Co, Trenton, N J, 1915

15 Judd, C H Psychological analysis of the fundamentals of arithmetic. *Suppl. Educ Monog*, 1927, No 32 Pp 122

16. Klapper, P The Teaching of Arithmetic New York. Appleton Century, 1934 (Pp 252-274)

17 Knight, F. B, & Behrens, M S The Learning of the 100 Addition Combinations and the 100 Subtraction Combinations New York· Longmans, Green, 1928. Pp 82.

18 Lazar, M Pupil's difficulties in learning fundamental operations in arithmetic. *Mod. Educ*, 1930, **3**, 10-11

19. MacLatchy, J H Counting and addition *Educ Res Bull.,* 1932, **11**, 96-100

20 ——— Another measure of the difficulty of addition combinations *Educ Res Bull*, 1933, **12**, 57-61

21. Olander, H T Transfer of learning in simple addition and subtraction. *Elem School J*, 1931, **31**, 358-69, 427-37

22 Osburn, W J. A study of the validity of Courtis and Studebaker practice tests in the fundamentals in arithmetic *J Educ Res*, 1923, **8**, 93-105

23. Phelps, C. L A study of errors in tests of adding ability *Elem. School Teach*, 1913, **14**, 29-39

24 Phillips, D E Number and its application *Ped Sem*, 1897, **5**, 254.

25. Ranschburg, P Zur phsiologischen and pathologischen Psycholgie der elementarem Rechenarten *Ztsch f Exper Pad*, 1908, **7**, 135-162; 1909, **9**, 251-263

26 ReBarker, H A A Study of the Simple Integral Processes of Arithmetic *Contrib Educ*, No 26, George Peabody College, 1926 Pp. 91.

27. Rock, R T, & Foran, T G A comparison of studies on the relative difficulty of the number combinations *Educ Res Bull*, Catholic Univ. America, 1930, **5**, No 8 Pp 26

28. Seaton, E T Correlation between *IB* and speed in addition *School,* 1926, 15, 166-167

29 Smith, J H Arithmetical combinations *Elem School J.,* 1921, **21**, 762-70

30 Wheeler, L R A study of the relative difficulty of a primary reading vocabulary *J Gent. Psychol*, 1938, **52**, 183-201.

31. ——— An experimental study of the value of informal methods in teaching primary reading *J Educ Res*, 1938, 335-347

32 Washburne, C, & Vogel, Mabel Are any number combinations inherently difficult? *J Educ Res*, 1928, **17**, 235-255.

33. Williams, C L, & Whitaker, R. L. Diagnosis of arithmetic difficulties *Elem School J*, 1937, **37**, 592-601.

State Teachers College
Johnson City, Tennessee

The Journal of Genetic Psychology, 1939, 54, 313-325

SOME RELATIONSHIPS BETWEEN STAIR CLIMBING AND PRONE PROGRESSION*

Yale Clinic of Child Development

Louise Bates Ames

A. Historical Note

Little study has been devoted to the problem of stair-climbing in the human infant As a rule early investigators (3, 11, 12) merely mention the age at which such behavior appears Tanner's statement (12, p. 347) that climbing when not repressed begins at about the same time as creeping, Fenton's (3, p 309) that at one year delight in climbing is very likely to be manifested; and Strang's (11, p. 106) that during the second year, usually at 18 months, the child learns to climb, are typical Shirley, also, discusses climbing onto objects and up and down stairs on hands and knees (10), but in no detail Trettein (13, p 34) in his excellent article on creeping gives the most detailed description of climbing behavior to be found in the early literature.

More recently we have the experimental study of Gesell and Thompson (7), a detailed investigation of the climbing behavior of identical twins The purpose of this experiment was to determine the effect of practice on such an activity as stair-climbing To this end, one twin (*T*), was trained in this behavior from 46 to 52 weeks. The second twin (*C*), was given a brief and delayed training period, from 53 to 55 weeks It was concluded that training and practice neither altered the total developmental complex nor hastened to any marked degree the manifestation of climbing behavior Descriptions of climbing behavior at ages from 46 to 80 weeks, and the number of frames required for movement of each foot are given, but the form aspects of behavior are not fully discussed.

The first normative study of climbing behavior is found in *Infant Behavior, Its Genesis and Growth* (8), also by Gesell and Thompson Climbing behavior as observed from 40 to 56 weeks is dis-

---

*Accepted for publication by Arnold Gesell of the Editorial Board, and received in the Editorial Office on June 9, 1938

cussed and norms are given. It is stated that by the age of 56 weeks, 53 per cent of the infants studied could surmount the four treads of the experimental staircase Reference is made to this source as the most detailed exposition of early climbing behavior. The present study will discuss patterned climbing behavior as seen from 56 weeks to three years.

B THE PRESENT INVESTIGATION

The purpose of the present study was, primarily, to compare the patterning of stair climbing behavior with the patterning of prone progression. Specifically, we wished to determine for climbing behavior, as has already been done for prone progression (2, p 444) the order in which the four limbs move, when they move in relation to each other, the exact duration of each limb movement, and the exact duration of the pauses between limb movements.

This investigation is based on a time analysis of the cinema records of the stair-climbing behavior of twelve subjects.[1] All subjects were observed from 52 to 80 weeks, and some to two and three years of age. Cinema records of their stair-climbing behavior were taken in the majority of cases at monthly intervals.

Since limb movement in patterned climbing is rapid, and since two limbs as a rule move simultaneously, use of the cinema was necessary to determine the exact timing of movements. Naturalistic observations were used only to supplement cinema records

Records of the prone behavior of each case being also available, the climbing behavior of each individual child was compared with the prone behavior seen contemporaneously Typical patterns of prone and climbing behavior were also compared. Since records were taken in each case over a period of months or even years, it was also possible to investigate the constancy of patterning in behavior, both as to form and tempo

Cinema records were taken in each case, of descending as well as of ascending behavior, and comparisons between ascent and descent were made, and differences noted.

---

[1] Eleven of the present subjects (Twins *T* and *C*, Girls *B*, *C*, *D*, and *G*-43, *G*-44, *G*-48, and Boys *B*-34, *B*-36 and *B*-38) are discussed also in references 6, 7, 8 and 9 which give detailed descriptions of these subjects and of the conditions under which they were studied.

C THE PATTERN OF STAIR CLIMBING BEHAVIOR COMPARED TO THAT OF PRONE PROGRESSION

Cinemanalysis of our records revealed a definite pattern in stair-climbing behavior, almost identical with that commonly observed in the normal *creeping* stage of prone progression

The typical stair-climbing pattern consists, as does the typical creeping pattern, of almost simultaneous forward movement of a contralateral pair of limbs, followed, after a short pause, by forward movement of the other pair of limbs. The chief difference between the climbing pattern and the prone pattern is that stair climbing often starts with a *single foot movement,* whereas creeping starts with movement of a *pair of limbs.*

In more detail—the typical *stair-climbing pattern* may be described as follows. It starts with placement of the left foot on the first stair. The left hand is then placed on the second stair, and immediately afterward, the right foot on the first stair There is a short pause, then right hand moves to the third stair and left foot to the second. After another short pause, the left hand and right foot move up again. And so it goes Temporal analysis of this behavior shows that left hand and right foot move almost simultaneously, as do right hand and left foot It is important to note that left hand and left foot do not move simultaneously, nor do right hand and right foot

Creeping usually begins with a forward hand movement Starting often simultaneously with such a forward hand movement is forward movement of the knee on the opposite side of the body The hand usually completes its forward movement just before the knee does so. There is a pause, and then a second hand and knee start forward, again hand completing forward movement just before the knee does so.

Time analyses of typical climbing and creeping patterns, with indications of the duration of each limb movement, are given in Tables 1 and 2

Creeping in the patterned fashion described above was observed in every one of a group of 20 cases previously studied (2) Contralateral limbs always moved simultaneously, though sometimes movement began with left hand and right knee instead of, as more commonly, with right hand and left knee

TABLE 1

QUANTITATIVE ANALYSIS OF A TYPICAL STAIR CLIMBING PATTERN[2]

Limb used	Movement starts	Movement is completed	Duration of movement (in seconds)
Left foot	0.00 seconds	0 93 seconds	93 seconds
Pause			.13 seconds
Left hand	1 06	1 62	56 seconds
Right foot	1.12	2 31	1 19 seconds
Pause			25 seconds
Right hand	2 56	3 12	56 seconds
Left foot	2 68	3 62	94 seconds

Total time in seconds for
whole pattern (four limbs) = 2.56 seconds
Pause between sets of limbs = 25 seconds

[2] This is a description of the behavior of Twin *T* at 56 weeks, about two months after her first successful climbing up a flight of four steps

TABLE 2

QUANTITATIVE ANALYSIS OF A TYPICAL CREEPING PATTERN[3]

Limb used	Movement starts	Movement is completed	Duration of movement (in seconds)
Right hand	0 00 seconds	0 45 seconds	45 seconds
Left knee	0 51	1 18 seconds	67 seconds
Pause			.22 seconds
Left hand	1 40	1.79	39 seconds
Right knee	1.51	2 35	84 seconds

Total time in seconds for
whole pattern (four limbs) = 2 35 seconds
Pause between pairs of limbs = 22 seconds

[3] This is a description of the behavior of Boy *B* at 48 weeks, shortly after he began to creep (2, pp 442-444)

Climbing in the patterned fashion described above was observed in every one of the present group of 12 cases, with but very minor variations. Eight children start climbing by lifting first a single foot, and then a contra-lateral pair of limbs as described; while four start with left foot and right hand, almost simultaneously. In two cases the pattern is reversed, movement starting with right foot instead of with left In one case the pattern is observed without the use of the right hand, the left foot moving alone; and in another case the left knee is used instead of the left foot It will be seen that these variations are minor and incidental, and do not significantly distort the typical pattern which has been described and which is basic to all.

The ages for the first appearance of this patterned climbing, that is, the age at which the child first climbed smoothly and without marked hesitation, up a flight of four or more stairs, are as follows: one child climbed successfully at 48 weeks; one at 52, six at 56 weeks, one at 60, two at 80 weeks; and one at 84. The mean age for this group, for first patterned climbing, is therefore *62 weeks.* In contrast, the mean age for first creeping is *44 weeks.* Thus on the average these infants crept about *18 weeks* before they climbed a flight of stairs.

D. Pattern of Prone Progression as Determining Type of Climbing Behavior

As might be expected, there is a marked tendency for stair climbing behavior to correspond to the mode of progression currently used on level surfaces. A child who creeps on hands and knees tends to climb stairs on hands and knees A child who creeps on hands, one knee and one foot, climbs stairs in that fashion A child who creeps on hands and feet, or one who walks, will climb upstairs on hands and feet, or if old enough, on feet only. In 11 children of the present group this correspondence was noted It is true, however, that a child who first climbs stairs while he is in the creeping stage, may persist in climbing upstairs on hands and knees long after he has begun to walk.

E Descent as Compared with Ascent

Gesell and Thompson (8, p 72) have commented

> The patterns of descent constitute another story, one which we have not investigated in detail The story is so distinctive that it cannot be explored by viewing a cinema of stair climbing in reverse For, from the standpoint of behavior mechanisms, descent is not simply a reversal of the process of ascent

Cinemanalysis reveals the chief differences between climbing upstairs and climbing down to be, differences in body orientation, in number of limbs used, and in order of limb movement

1 *Differences in Body Orientation*

In the case of Girl *C*, as will be noted from the listing below, reversal of the cinema record of ascent would give the correct order of limb movement for descent. It will be seen, however, that the child starts upstairs with the left knee; down, with the right hand. Furthermore, in ascent the left knee only (not the left foot) is used, in descent, the left foot contacts each stair before the knee comes to rest on that stair. Also the body orientation differs in the two instances, being determined in each instance by the goal in sight.

TABLE 3
ASCENT AS COMPARED WITH DESCENT
Girl *C*, 61 weeks

	Top	Top	
	Right hand	Right hand	
	Right foot	Right foot	
		(left foot)	
	Left hand[4]	Left hand	*Read down*
Read up	Left knee	Left knee	
	Right hand	Right hand	
	Right foot	Right foot	
		(left foot)	
	Left hand	Left hand	
	Left knee	Left knee	
	ASCENT	*DESCENT*	

---

[4] The spaces in the ASCENT column do not indicate pauses, but are left that the similarity of limb order in the two operations may be made clear

2. *Number of Limbs Used*

Infants as a rule tend to use fewer limbs in patterned descent than in ascent Often one or both hands are not used in descent

For example, Twin *T* ascends the stairs at three years, as at all ages from 52 weeks on, by use of hands and feet, while in descent she uses only her feet and right hand The limb order in descent can be obtained merely by reversing the order seen in ascent (and omitting the limb not used). Twin *C* uses neither hand as she descends the stairs at three years Moreover, she starts to climb down with the foot opposite to that with which she starts to climb up

3. *Another Common Type*

The child sits on one stair facing down, with both feet on the stair immediately below, and then lowers the feet in turn to the second stair below. For instance, Child *G*-44 (80 weeks) sits on one stair and places her right foot, then her left, on the second stair below. Many children simply face the foot of the stairs and walk down, two feet to a step This behavior often appears while the child is still creeping up the stairs, using all four limbs, as in the case of Girl *B* (80 weeks).

Patterned descent as a rule first occurs within a few days of patterned ascent, at least in the present group of cases.

F. CONSTANCY OF PATTERNING

Although it is obvious that practice and increasing age of the climber will to some extent alter climbing behavior, we wish to emphasize here the rather remarkable constancy both in pattern and speed which can as a rule be observed in the climbing behavior of any one infant from age to age

It is significant that the child often climbs one stair, or several stairs, the first time he is placed before a flight of stairs (8, p. 70, 7, p 60) Or, perhaps more commonly, he will first ascend a complete flight of stairs, with patterned rhythmic movement, only a week or ten days after he has first surmounted the first tread. Then, from the time when this patterned behavior is first exhibited, it continues, in the same pattern and with but slight increase in speed, often for over a year, until the child eventually walks upstairs, two feet and finally one foot to a stair.

The case of Girl *C*, a child studied daily for over two years, has been selected as a good illustration of the early appearance and marked constancy of patterned behavior in stair climbing. Girl *C* first climbed up one stair on the *416th day (59 weeks),* using left knee, left hand, right knee, right hand, in that order. On the *418th day* she climbed one stair, and on the *426th day,* two stairs, using left knee, left hand, right hand, right foot By the *427th day* she climbed eight steps in the pattern of left knee, left hand and right foot, right hand and left foot. This pattern persisted until the *627th day* (20 months) On the *627th day* she walked upstairs, holding the rail with her left hand and lifting her left foot, then her right, to each step.

The same quick onset and constancy of pattern is observed in the descent behavior of this child On the *426th day* (60 weeks) she climbed down one stair. Facing the top, she lowered right knee, then left down beside it On the *430th day* she climbed down six steps in the manner which was observed thereafter till the 20th month, that is she half faced the top and lowered right hand, right foot, left hand, left knee in that order, left foot contacting each stair just before the knee came to rest *After 20 months* she walked downstairs, two feet to a step, when held by the hand.

TABLE 4

TEMPORAL DESCRIPTION OF STAIR-CLIMBING BEHAVIOR OF TWIN T

Limb moving	Duration of Movement in Seconds 52 weeks	56 weeks	80 weeks	92 weeks	2 years	3 years
Left foot	1 06	93	75	68	62	.62
	(.25)	(12)	(12)	(.12)	(12)	(−25)
Left hand	.81	56	56	56	.56	.56
	(−18)	(−50)	(−31)	(−18)	(−.31)	(−18)
Right foot	1.31	1 25	75	75	.68	68
	(.25)	(.25)	(12)	(18)	(12)	(−18)
Right hand	75	56	56	56	50	50
	(−18)	(−43)	(−12)	(−31)	(−18)	(−18)
Left foot	1 06	.93	75	68	62	62
Whole pattern	3 62	2 62	2 31	2 25	2 00	1 81
Wait between sets of limbs	(.25)	(.12)	(12)	(12)	(12)	(−25)
	(25)	(25)	(.12)	(.25)	(12)	(−18)

TABLE 5

TEMPORAL DESCRIPTION OF STAIR-CLIMBING BEHAVIOR OF TWIN C

Limb moving	Duration of Movement in Seconds 52 weeks	56 weeks	80 weeks	92 weeks	2 years	3 years
Left foot		.93	87	56	56	.62
		(25)	(37)	(18)	(.12)	(−.06)
Left hand		62	62	.50	50	56
		(−.25)	(−.31)	(−06)	(−31)	(−06)
Right foot		1 00	1 06	56	56	62
		(.12)	(37)	(06)	(06)	(−18)
Right hand		68	62	50	43	50
		(−25)	(−12)	(−12)	(−18)	(−18)
Left foot		.93	87	.56	56	.62
Whole pattern		2.87	3.12	2 00	1 62	1 87
Wait between sets of limbs		(25)	(.37)	(37)	(12)	(−06)
		(.12)	(37)	(06)	(06)	(−18)

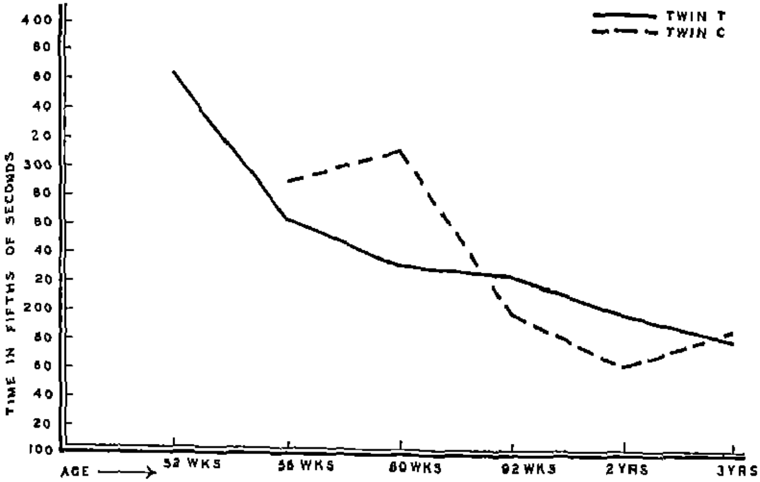


FIGURE 1
Time for Whole Pattern
T and *C* Compared

A detailed quantitative study of the climbing behavior of two identical twins, *T* and *C*, also illustrates the comparatively small changes which take place in patterned climbing from age to age Summary Tables 3 and 4 and graphs (Figures 1 and 2) present their behavior at all ages when patterned climbing was observed These are based on figures obtained by averaging all cinema records of climbing at each age. These records varied in length from 12 to 88 seconds of recorded behavior at an age, and represented from two to eight ascensions of an entire flight of five steps. In ascending even a five-step flight of stairs, the infant of necessity exhibits the complete four limb pattern several times Thus the averages given in the tables appear to be representative.

Figures given indicate the number of seconds' duration, on the average, of movement for the limb in question Figures in parentheses indicate time in seconds between lowering to the step of one limb, and lifting of the following limb. A minus quantity means that a second limb lifts before a first has come to rest. A plus quantity means that there is a pause between the lowering of one limb and the lifting of the next It is evident that in each

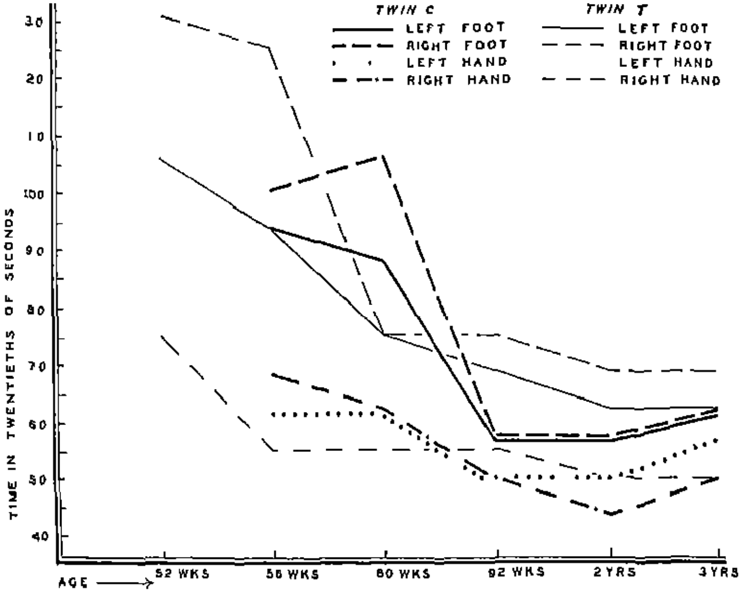


FIGURE 2
TIME FOR INDIVIDUAL LIMB MOVEMENTS*
T and *C* Compared
*When two lines are drawn in one interval they may be interpreted as being identical

instance one hand and the contralateral foot are moving simultaneously, but never a hand and foot on the same side of the body.

In general it is clear that in the case of each twin, *there is a marked temporal constancy of behavior from 52 weeks to 3 years.* The greatest decrease in time (from 1 31 seconds to 68 seconds) occurs in the case of Twin *T*'s right foot. There is considerably less decrease in speed for most limb movements; for example, a decrease on only 31 seconds in two years for Twin *C*'s left foot, of only 06 seconds in two years for *C*'s left hand. In fact, in the case of Twin *C*, the decrease in time for the whole four limb pattern over a two year period is only *one second.*

It is interesting to note that there is an almost steady, though small, decrease in the time for movement of each limb with increasing age, as well as a tendency for the time for the whole pattern to decrease. An exception to this trend is the fact that the values of Twin *C* at 3 years are slightly larger than those at 2 years, because at 3 years she used one foot to a stair instead of two feet to a stair. This means that she had to take a longer step with each foot. Although Twin *T* also climbed with only one foot to a stair at 3 years, she showed no increase in time of movement at this age. Furthermore, at 80 weeks, Twin *C* moves more slowly than at 56 weeks, apparently because she uses her right knee instead of her right foot

Not only does the absolute time for limb movement remain constant over a period of two years, but the relative time as well For example, in the case of Twin *T,* the left foot always moved more quickly than did the right, even though the time difference was small Her right hand moved more quickly or at exactly the same speed as her left hand at all ages—never more slowly. The pause between movement of sets of limbs is also relatively constant, and it will be noted that the pause when not exactly the same, is always longer between the same limbs: for *T,* between right foot and right hand; for *C,* between left foot and left hand.

Comparison of the tables further reveals many striking similarities in the timing of the behavior of the two infants For example, the left foot of each child at 56 weeks takes .93 seconds to complete movement; at 3 years, 62 seconds. At three years the left feet, left hands, and right hands of the twins move at exactly the same speed.

It is interesting to note, in view of the fact that Twin *T* was trained in climbing from 46 to 52 weeks (7) and Twin *C* was not trained till she was 53 weeks and then only for two weeks, that at only two of the five age levels when data for both are available, is Twin *T*'s climbing appreciably speedier than that of Twin *C*. The trained twin is speedier at 56 and 80 weeks, but after that the untrained twin catches up and climbs the more quickly However, there is at all ages an almost negligible difference of less than one second between the two children in the time required for the whole four limb pattern

G CONCLUSIONS

Stair climbing in the human infant is definitely patterned behavior, almost identical to that commonly observed in the normal *creeping* stage of prone progression.

Typically, the left foot moves first; then left hand and right foot, almost simultaneously, then, after a short pause, right hand and left foot This basic pattern holds from child to child though some start to climb with the right foot, or with one hand, instead of with the left foot As in prone progression, contralateral limbs move simultaneously

Climbing behavior as a rule agrees with the type of prone progression currently manifested. A child who creeps on hands, one knee and one foot will climb, using these limbs, a child who walks, or creeps on hands and feet, will climb using feet rather than knees.

Patterns of descent differ from those of ascent in several particulars· in body orientation, in number of limbs used, in order of limb movement and in body posture

Although climbing upstairs on hands and knees gradually gives way to walking upstairs, two feet and finally one to a stair, as the child grows older, there is a remarkable constancy in the climbing behavior of any given (normal) infant throughout the second year of life both as to pattern and speed of movement. Quantitative tables compiled from cinemanalysis bear out this statement

In the present group of cases, patterned creeping behavior first appeared at a mean age of 44 weeks, patterned ascent of stairs at a mean age of 62 weeks, and patterned descent a few days later

REFERENCES

1 AMES, LOUISE B. Development of motor coordination in one child from birth to two years Unpublished M A Thesis, 1933 University of Maine Library.

2 ——— The sequential patterning of prone progression in the human infant *Genet Psychol Monog*, 1937, **19**, 409-460

3 FENTON, J C A Practical Psychology of Babyhood New York: Houghton Mifflin, 1925. Pp xvi+348

4 GESELL, A. Prone progression in the human infant (35 mm. film)

5. ——— Normative film Stair climbing behavior from 40-56 weeks (16 mm)

6 GESELL, A *et al.* An Atlas of Infant Behavior New Haven Yale Univ Press, 1934 (2 vols) Pp 922

7. GESELL, A., & THOMPSON, HELEN Learning and growth in identical infant twins an experimental study by the method of co-twin control. *Genet Psychol Monog*, 1929, **6**, 1-124

8 ——— Infant Behavior. Its Genesis and Growth New York. McGraw-Hill, 1934 Pp viii+343.

9 ——— The Psychology of Early Growth Including norms of infant behavior and a method of genetic analysis New York Macmillan, 1938 Pp ix+290

10 SHIRLEY, MARY M The First Two Years A study of twenty-five babies (Vol I) Minneapolis Univ Minnesota Press, 1931 Pp xv+227

11 STRANG, RUTH An Introduction to Child Study New York Macmillan, 1930

12 TANNER, AMY. The Child. New York Rand, McNally, 1915

13 TRETTIEN, A W Creeping and walking *Amer J Psychol*, 1900, **12**, 1-57

Clinic of Child Development
Yale University
14 Davenport Avenue
New Haven, Connecticut

The Journal of Genetic Psychology, 1939, 54, 327-336.

THE INTELLIGENCE OF A NEGRO HIGH SCHOOL POPULATION IN A NORTHERN CITY*

Chicago Board of Education

ALBERT SIDNEY BECKHAM[1]

This is a study of the intelligence of Negro boys and girls in the Du Sable High School in Chicago. This is a city school with more than three thousand pupils

Suitable samplings were obtained from the first, second, third, and fourth years of the school These pupils were given the Henmon-Nelson Tests of Mental Ability, High School Examination. This test is suitable for grades seven to twelve It takes 30 minutes to administer this test. The coefficient of reliability of this test is 882 (1)

The Du Sable High School is a rather cosmopolitan one It is varied in reference to economic and social status. These pupils represent the following types of families according to the Taussig Scale of economic-social status:

1	Unskilled (families)	20 90%
2	Semi-skilled (families)	25 00%
3	Skilled (families)	36 05%
4	Semi-professional (families)	12.07%
5	Professional (families)	5 95%
	Total	99.97%

The intelligence of each year of the high school has been studied separately Each age group within the class has also been studied

What about the distribution of intelligence in this group? Is it

---

*Received in the Editorial Office on June 17, 1938

[1] For this study, I owe much to Mr C C Willard, Principal of the Du Sable High School, for many practical suggestions and much of the material that made the study possible

The encouragement of Dr Grace Munson, Director of the Bureau of Child Study, and Dr William H Johnson, Superintendent of Schools, has been most beneficial I also take this opportunity to express my appreciation to my assistants, Mr. Fred Carter, Mr Robert Lawrence, and Miss Georgia Williams homogeneous or is it a skewed distribution? The relationship of this distribution and Freeman's findings is interesting

> In an actual experimental situation in which a study is conducted with a sampling of any given population, we do not expect to find a perfect normal distribution, known also as the Gaussian Curve The occurrence of such would be a matter of chance *However, we do expect to find a reasonably good* approximation to the theoretical curve (1, p 394).

The intelligence of the high school freshmen will be presented in Table 1

TABLE 1

Shows the Intelligence of 284 High School Freshmen at Various Age Levels According to the Henmon-Nelson Test of Intelligence, Also the Per Cent at Each Age Level, and Grade Expectancy According to Age

Age group	Per cent	*IQ* Range	Mean *IQ*	*SD*	Grade expectancy according to age
12	0.69	102—144	123	——	6.6
13	9.85	83—143	105.7	15.4	7.8
14	27.11	72—125	94.4	11.8	8.5
15	31.33	71—118	86.0	8.4	
16	18.30	68—110	83.3	17.5	Second year High School
17	8.09	68—110	86.3	9.6	Third year High School
18	4.22	62—96	81.5	9.8	Senior year High School
19	0.38	———	87.0	——	High School Graduate
Total	99.97	62—144	93.4	12.8	
CA Range		12–19			
Mean age		14.9			
No. of pupils		284			

The intelligence of High School freshmen is of widening interest. This is as it should be, because it is important indeed to select the proper type of pupil at the beginning if we are carefully concerned with the finished product of the high school

We had more than a thousand freshmen to select from The method of selection was to take every tenth name in the various divisions. (The high school population is divided into a number of divisions. Each division is numbered and assigned to a specific teacher) Thus we see from Table 1 that the age range of these 284 freshmen is from 12 to 19

It is readily seen then, that the freshman class had an age range of seven years. Ages 12 and 19 were represented by a very few pupils The mode for these 284 freshmen was age 15 At this age we have 31 33 per cent of the class The majority of the class centers around ages 14, 15, and 16 Above age 16 we have 12 69 per cent Below age 14 we have 10 54 per cent The mean intelligence of the twelve year old members of the freshman class was 123 whereas the mean for the nineteen year old freshmen was 87.0 The mean *IQ* for the 284 freshmen of all ages was 93.4.

The freshman class according to Henmon-Nelson Test was one of average intelligence, with an *IQ* range from 62 to 144 The standard deviation for the class was 12 8 The mean age for these 284 freshmen was 14 9.

The mean for the *IQ*'s at the various age levels in the freshman population ranged from 81.5 for members of the class 18 years of age to 123 for year twelve

According to normal grade expectancy 37.65 per cent of the freshmen are accelerated and 30 99 per cent are retarded in educational progress The per cent placed normal according to grade expectancy was 31 33

If a pupil makes normal grade progress he enters the high school around his fifteenth year He should graduate at eighteen

Table 1 reveals that the freshman class is average both in respect to age and intelligence The younger members of the class were on the average not only accelerated according to grade, but more intelligent as well. The older members of the class were, on the average, retarded according to grade expectancy, and also less intelligent The class as a whole was average.

Table 2 shows the intelligence of second year high school pupils These boys and girls present a wide range in chronological ages. They were selected in the same manner as the freshmen Each tenth name was chosen from the several second year divisions. There were 264 pupils examined in the second year. The age range for the second year pupils was from 13 to 20 The youngest pupil had just passed her thirteenth birthday and the oldest was 20 years and 4 months. There was an age range of seven years among these 264 pupils. Most of these pupils clustered around the years fifteen, six-

TABLE 2

Shows the Intelligence of 264 Second Year High School Pupils at Various Age Levels According to Henmon-Nelson Test of Intelligence, Also the Per Cent at Each Age Level, and Grade Expectancy According to Age

Age group	Per cent	*IQ* Range	Mean *IQ*	*SD*	Grade expectancy according to age
13	0.49	———	100	——	7.8
14	4.90	82—125	100.5	——	8.5
15	24.50	72—114	90	10.0	9.7
16	26.42	77—117	91	8.16	Second year High School
17	25.00	67—111	90.4	8.12	Third year High School
18	11.76	71—122	85.8	9.6	Senior year High School
19	6.37	68—117	90.5	——	High School graduate
20	0.49	———	74	——	High School graduate
Total	99.93	67—125	91.5	8.97	
CA Range		13–20			
Mean age		15.6			
No. of pupils		264			

teen, and seventeen. Around these three years we find 75.92 per cent.

Below age 15 we find 5.39 per cent of the class examined. About age 17 we find 18.62 per cent of the class.

The question may be asked—Is the higher intelligence found among the younger or older members of the class? Let us see. Pupils below 15 years would be considered young for second year in high school. The mean *IQ* for pupils of age 13 and 14 was 100.

Age 18, 19, and 20 would be considered old for second year in high school. The combined mean *IQ* for age 18, 19, and 20 was 83.4. The combined mean *IQ* for age 15, 16, and 17 was 90.4.

These figures reveal that the younger pupils were on the whole, average in intelligence, while the older ones were dull.

The mean *IQ* for the 264 pupils examined was 91.5, with a standard deviation of 8.97.

The dullest pupil examined in the second year class had an *IQ* of 67, whereas the brightest pupil had an *IQ* of 125.

These pupils representing a class would be considered average in intelligence according to the Henmon-Nelson Test of Intelligence.

The mean age of the class was 15 years and 6 months.

If, according to Table 2, a 16-year old pupil should be in the second year in high school, and there are 43 6 per cent of these pupils who are older than 16, it is clearly revealed that this per cent of these pupils is retarded according to grade expectancy.

About 30 per cent of these pupils were accelerated according to grade expectancy, and about 26 per cent were properly placed in school according to age.

The sophomore, like the freshman class was one of normal or average intelligence

From the high school juniors we selected 187 pupils This is a fair sampling from a population of 356. Selections were made in the same way as was done in the case of the freshmen and second year classes.

No pupil in the class was younger than 14 years The oldest person in the class was 22 years of age The age range for these 187 juniors was eight years, that is from 14 to 22

The mean chronological age was 16 9 Age 14 and 15 would be considered young for high school juniors Likewise 19 and above would be considered old for the junior class

The mean *IQ* for ages 14 and 15 combined was 103 7 At these combined age levels we have 12 29 per cent of the 187 juniors who represent the class. The mean *IQ* for age 19 and those above 19 combined, was 89 The mean *IQ* for the whole class of 187 was 92 25

In other words, the juniors were a class of pupils representing average intelligence according to Henmon-Nelson Tests of Intelligence The *IQ* range was from 69 to 125. The standard deviation was 8 27

The mode for the juniors was age 17 At this age level we find 31 55 per cent, with an *IQ* range from 73 to 112

Table 3 reveals that 36 35 per cent of the juniors were accelerated in grade accomplishment according to their chronological age About 32 per cent were properly placed according to their chronological age About 32 per cent of the class was retarded according to age The over age juniors were on the average less intelligent than the younger members of the class On the average the class possessed normal intelligence

Table 4 reveals the intelligence of high school seniors The selec-

TABLE 3

Presents the Intelligence of 187 Third Year High School Pupils at Various Age Levels, According to Henmon-Nelson Test of Intelligence, with Per Cent at Each Age Level, and Grade Expectancy According to Age

Age group	Per cent	*IQ* Range	Mean *IQ*	*SD*	Grade expectancy according to age
14	2.13	96—125	114	——	8.5
15	10.16	81—112	93.4	8.36	9.7
16	24.06	72—115	93.6	8.12	Second year High School
17	31.55	73—112	84.4	8.8	Third year High School
18	22.45	73—107	85.6	7.8	Senior year High School
19	6.41	69—123	90.0	——	High School graduate
20	2.67	70—100	85.0	——	High School graduate
Over 20	0.53	———	92.0	——	High School graduate
Total	99.96	69—125	92.25	8.27	
CA range		14–22			
Mean age		16.9			
No of pupils		187			

TABLE 4

Presenting the Intelligence of 177 High School Seniors, at Various Age Levels, According to Henmon-Nelson Test of Intelligence, Per Cent at Each Age Level and Grade Expectancy According to Age

Age group	Per cent	*IQ* Range	Mean *IQ*	*SD*	Grade expectancy according to age
15	2.41	99—105	97.00	——	9.5
16	18.33	86—127	100.8	13.5	Second year High School
17	31.66	73—124	97.2	12.28	Third year High School
18	26.66	71—118	93.6	13.1	Senior year High School
19	15.83	71—116	85.0	10.8	High School graduate
20	2.49	93—112	108.0	——	High School graduate
Over 20	2.49	77—83	79.0	——	High School graduate
Total	99.87	71—127	94.4	11.42	
CA range		15–22			
Mean age		17.5			
No of pupils		177			

tion here was somewhat different. Seniors represent a selective process in the high school population.

Every fifth name from the senior divisions was chosen. There were 177 persons selected to represent the senior class.

The youngest senior was 15 years of age and the oldest was 22 years and 3 months. The age range for the seniors was seven years. The mean chronological age was 17.5.

For the senior class age 15 and 16 are considered young. At these ages combined, we find 20.74 per cent of the 177 pupils studied. The combined mean *IQ* for these two ages was 98.5. In other words, the younger students, those who were 15 and 16 years of age, are considered as of average intelligence as a group.

The older students, those who are 20 and over, have a mean combined *IQ* of 89.5.

The mean *IQ* for the seniors, representing all ages from 15 to 22 was 94.4, with a standard deviation of 11.42. The *IQ* range for the whole group was from 71 to 127.

The mode for the seniors was age 17. Here we find 31.66 per cent of the group. The *IQ* range at this age level was from 73 to 124.

The seniors represent a group of students with average intelligence. Each age group among the seniors was average except age 19 and those over 20 years of age.

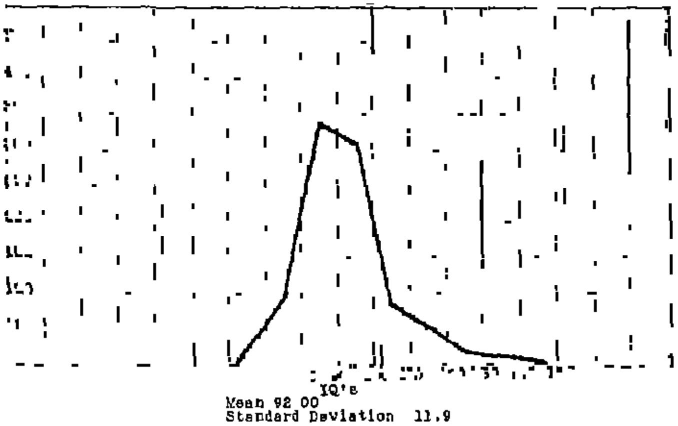


FIGURE 1
THE DISTRIBUTION OF *IQ*s OF A NEGRO HIGH SCHOOL POPULATION

In the senior class we have 52 per cent who are accelerated in grade placement according to age About 27 per cent of these boys and girls were properly placed in grade according to age The table reveals that about 21 per cent of the seniors were retarded according to their chronological age

Figure 1 reveals the curved distribution of 912 high school pupils, representing each of the four years, freshman, sophomore, junior and senior The mean *IQ* for the school was 92. The Standard Deviation was 11.9

Figure 2 gives a diagram of the chronological age for each of the

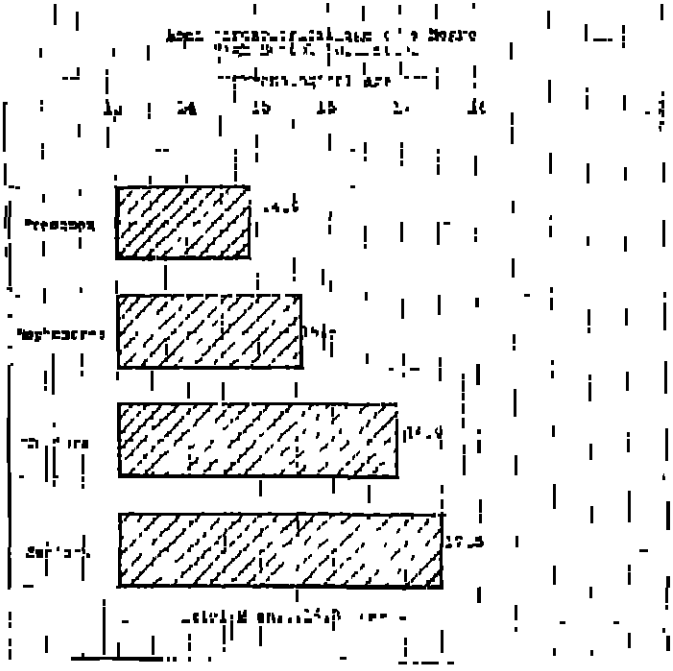


FIGURE 2

four classes, freshman, sophomore, junior, and senior The mean age for the school was about 16 years

Table 5 gives the distribution of *IQ*s for a sampling of the freshman, second year, third year, and senior classes.

TABLE 5

DISTRIBUTION OF *IQ*s ON THE HENMON-NELSON TEST OF INTELLIGENCE OF A NEGRO HIGH SCHOOL POPULATION, SHOWN IN PERCENTAGES

Class	No	Per cent	*IQ* Distribution 60	70	80	90	100	110	100	130	140
Freshman	284	99 96	3 52	11 32	33 09	37 67	11 86	10 16	1 69		
Sophomore	264	99 99	1 13	11 36	36 12	32 57	7 74	4 87	0 70	0 70	0 35
Junior	187	99 96	0 53	7 48	40 71	33 08	12 36	4 54	1 51	0 37	
Senior	177	99 96		7 90	33 58	34 77	11 76	5 34	1 06		
Total No	912		14	93	336	308	95	51	11	3	1
Total per cent		99 94	1 53	10 19	36 84	33 78	10 41	5 39	1 20	0 32	0 10

Class	Mean *IQ*	*IQ* Range	*SD*	Mean *CA*
Freshmen	93 4	62 to 144	12 8	14 9
Sophomore	91 5	67 to 125	8 97	15 6
Junior	92 25	69 to 125	8 27	16 9
Senior	94 4	71 to 127	11 42	17 5
Total No	912			
Total Mean *IQ*	92 0			
Total *SD*	11 9			

According to Terman Classification of *IQ*s we have 1.53 per cent with defective intelligence. We have 10 19 per cent who were classified as border line. About 37 per cent classified as dull. Those who classified as having average intelligence, that is to say, with *IQ*s between 90 and 109, were 44 19 per cent. For pupils representing superior intelligence there were 5 39 per cent. Fourteen pupils or about two per cent had *IQ*s above 120 One pupil had an *IQ* of 144.

As to the ages of these pupils the table reveals that the average age for the freshmen was 14 years and 9 months, for the second year pupils 15 years and 6 months; for the juniors the average age was 16 years and 9 months, and for the seniors the average age was 17 years and 5 months

It must be stated here that these tests were made nearer the beginning than the end of the school year The ages here reported would be nearer the beginning than the end of any specific year class.

REFERENCES

1 SKINNER, C E. Educational Psychology New York. Appleton, 1926

4900 Wabash Avenue
Chicago, Illinois

The Journal of Genetic Psychology, 1939, 54, 337-352

SOCIAL RELATIONS AND PERSONALITY PATTERNING IN PRESCHOOL CHILDREN*[1]

New York State Bureau of Child Development and Parent Education

EUGENE L. HOROWITZ AND RANDOLPH B. SMITH

A.

Objective studies of personality patterns in young children have been handicapped, in general, by the inability to secure data from a sufficiently large sample to warrant statistical treatment. The opportunities for research offered by the federally assisted public nursery schools provided the present writers' data—collected under similar conditions, in a number of schools, from a population homogeneous with respect to general socio-economic status. In thirteen schools in eight different communities in New York State (outside of New York City) coöperation of teachers was secured and comparable data collected. These schools were under local supervision but represented a relatively unified philosophy of education for young children and had approximately the same school equipment, though the physical lay-outs differed. Ethnic backgrounds of the children differed too.

Starting in April and May, 1936, records were taken of children's behavior on consecutive school days. Children were observed for half minute periods during free play, indoors and out of doors, and the material collected in 20 such records for each child was organized as the basic data for this report. Absences interfered with a perfect record of behavior on *successive* days as did week-ends and holidays.

---

*Recommended for publication by Ruth Andrus, and received in the Editorial Office on June 22, 1938.

[1] This is an abbreviated report of a study undertaken as part of the Nursery School and Parent Education Program under the supervision of the Child Development and Parent Education Bureau of the New York State Department of Education, Dr. Ruth Andrus, Chief, funded by the *TERA* and the Federal Works Progress Administration of New York State. The data here presented were collected as part of a project conceived by Randolph B. Smith (now Executive Secretary, Cooperative School for Student Teachers, New York City) who supervised the gathering of the material. The analysis was undertaken under the direction of Eugene L. Horowitz (State Supervisor in the program) who is chiefly responsible for its preparation in the present form. A more complete report is on file at the Child Development and Parent Education Bureau at Albany.

The selection of schools for participation in this investigation was restricted to schools which had been in operation for some time, i.e., schools in which teachers were sufficiently familiar with the normal work of a nursery school to be able to afford time for the record-taking. The particular schools were chosen on the recommendation of the state and regional supervisors, undertaking to select schools *in which the teachers were likely* to be most interested and conscientious in participating in the study.

One week's training of the teachers who were to participate in the study was undertaken, the research proper beginning the following week. This training was carried on by the research director, in the individual schools. This training period involved a Monday afternoon staff meeting with the entire nursery school staff to explain the purpose and procedures of the study and to provide general orientation. Tuesday through Friday the research director worked with the individual teachers taking simultaneous records with them so that misinterpretations or other difficulties in procedure could be cleared up on the spot. Records were taken independently and then checked with full discussion of any differences in interpretation or divergence in classification. In general, for the duration of the study, all the records in each school were taken by only one person, though occasionally it was necessary to substitute another observer. In schools in which the group of children was divided on an age basis for the routine of the day with different activities for the different groups, the teacher most familiar with each group took the records and they were combined for the school.

A standard form for time-sample records was set up to permit one 30-second observation of each child each day. All observations were to be taken during free-play periods when the child was less subject to adult controls. The names of the children were listed alphabetically on the record blank and the teacher making the observation was instructed to proceed directly to the observation of each child in turn as soon as the preceding observation and recording was completed, regardless of the activity of the child, just so long as he was not obviously subject to adult interference. This was to provide for a strictly chance sequence of observations without possibility of selecting particular types of behavior. The alphabetical order of the observations was rotated by one name at a time each day—the entire list "moving up one" each day This was in order to avoid constant errors such as would be introduced by regularly observing the first part of the alphabet early in the nursery school day and last part regularly later in the day Time was taken not with stop-watches but with watches with minute hands, the time of beginning the observation being recorded on the blank to the nearest minute. (The purpose of this time record was to provide for possible analysis of the data in terms of records taken at different times in the nursery school day.) During the 30-second period the observer concentrated entirely on observing, doing no recording until after the observation period was completed The behavior was then immediately analyzed and recorded in the appropriate categories of the blank before going on to the observation of the next child on the list

Entries in the different categories were made with numbers rather than with simple tallies in order to make possible subsequent analysis of possible patterns in the sequence of behavior. The first type of behavior which occurred in a given observation interval was recorded in the appropriate cell by entering a number one If other types of behavior occurred in the same observation interval for a given child these were recorded by the numbers 2, 3, 4, etc, according to the actual sequence of different types of behavior

The definitions of the categories of the time-sample observation blank were set up fairly explicitly, though in the training of the teachers special emphasis was placed upon agreement in understanding of concepts rather than the formalized definitions

Under the general heading "Type of Play" three different kinds of solitary activity were distinguished These three categories of solitary play were considered to represent three different levels of approach toward group activity Parallel activity and group activity completed the general sequence Unfriendly, combatant behavior was studied with attention to the question whether the child being observed occupied the rôle of recipient of aggression Friendly relations were similarly studied, did the child initiate or join the activity?

The data thus collected permit numerous analyses. For the present report we have limited ourselves to the study of the interrelation of the observed forms of behavior The differences in frequency of manifestation of the various forms of behavior considered in the different schools indicated that the sampling could not accurately be

TABLE 1

Coefficients of Rank Difference Correlations Among the Categories of Behavior in Four Schools, *B*, *H*, *I*, *K*

Key

Schools		No. of cases	
I	*H*	23	24
K	*B*	23	20

	2		3		4		5		6		7		8	
1	49	36	48	.51	22	24	—.14	—58	—05	07	—36	—15	33	33
	22	31	.50	32	05	—03	10	—66	—18	11	—03	31	11	18
2			.65	36	34	42	—24	—56	.17	20	—29	12	27	27
			—05	25	40	—19	—19	.02	.41	17	32	.48	39	14
3					—02	45	19	—35	32	29	—23	—35	31	57
					—03	—21	.26	—14	—17	—32	15	45	15	20
4							—07	43	29	50	22	29	25	62
							06	—03	.34	49	.28	39	40	27
5									05	16	28	09	53	—13
									—11	01	16	25	10	30
6											.08	15	.51	26
											69	32	66	42
7													12	13
													66	55
8														
9														
10														
11														
12														

TABLE 1 (*continued*)

	9		10		11		12		13		14		15		16	
1	09	.59	.09	50	43	49	01	54	53	35	—45	—35	—35	—15	—87	—37
	24	19	—13	44	40	45	05	15	30	38	06	—25	—18	18	—68	—62
2	29	49	11	39	56	58	02	49	49	55	—01	.25	—20	—16	—31	—37
	26	33	18	24	44	59	17	.32	06	50	.23	—15	30	20	—03	—79
3	15	60	29	40	55	.52	—06	56	41	50	04	—21	—05	—34	—39	—38
	17	—57	—22	35	38	92	10	05	20	30	14	—15	—08	25	—43	—21
4	65	69	20	75	66	89	—06	82	68	69	13	48	06	35	07	49
	38	06	.05	17	53	45	58	—06	32	38	33	.24	07	28	—08	05
5	23	.04	28	36	36	49	37	34	46	08	04	47	45	28	—12	72
	17	37	—04	24	41	32	36	32	43	33	—27	22	—17	08	—50	36
6	30	35	23	50	54	59	24	47	55	43	.40	16	12	—09	07	00
	61	42	63	42	66	49	53	29	44	34	29	—14	58	02	48	09
7	44	29	.55	54	.46	56	07	50	46	47	34	94	.89	40	54	—94
	51	55	.40	70	57	77	44	47	66	55	—02	51	40	60	15	.30
8	41	62	35	59	65	66	46	71	.70	65	42	—86	28	35	16	—90
	.46	85	34	70	72	63	62	78	55	70	37	39	28	14	.12	44
9			47	83	77	.84	,00	90	66	76	.39	12	26	34	46	13
			26	70	70	63	41	.95	62	70	36	21	68	27	49	29
10					65	89	41	.95	65	69	45	20	65	44	46	.41
					57	77	.57	62	19	55	.16	.43	39	55	31	.40
11							44	94	09	.83	42	.52	50	54	.52	52
							78	56	61	63	51	.35	50	48	53	30
12									44	.77	19	35	44	47	—04	38
									44	62	41	25	20	31	25	31
13											46	13	50	.29	38	07
											19	43	.33	27	05	05
14,													.43	23	54	59
													22	10	20	34
15															58	22
															65	01

Categories:

1	Sedentary	9	Retaliator
2.	Moving around	10	Verbal resistance
3	Attentive interest in others.	11	Screaming, stamping
4	Parallel activity	12	Appeal to adult
5	Group activity	13	Running away
6.	Physical violence	14	Offers invitation.
7	Threats, teasing, bullying	15	Accepts invitation
8	Passive recipient	16	Constructive social use of materials

TABLE 2

INTERCORRELATIONS AMONG THE CATEGORIES OF BEHAVIOR IN 5 SCHOOLS, *C, G, L, E, M*

Key

Schools		No of cases	
C	*E*	19	19
G	*M*	13	13
L		13	

	2		3		4		5		6		7		8		9	
1 Sedentary	—32	05	.13	02	—.32	—19	—72	—19	—12	10	—18	—.16	15	14	—28	—43
	30	46	68	54	08	—03	—73	—79	—39	—06	—49	—17	.23	—60	—26	—45
	48		29		—35		64		35		—23		.11		—.42	
2 Moving around			11	—15	—04	—09	—09	07	.11	—02	—02	27	—30	.04	—14	15
			24	.48	53	53	—.28	—56	—09	07	—16	46	—07	43	—35	.38
			21		30		65		42		11		16		—30	
3 Interest in others					—52	02	10	—.32	—25	—19	—30	—19	16	—08	—37	—11
					32	29	—87	—22	—43	15	—.49	00	—15	19	—.28	00
					—30		—31		—06		57		40		—.63	
4 Parallel activity							.24	03	06	.01	.14	—13	15	02	09	—06
							—.47	18	—48	44	—46	.63	—49	—11	—61	—.08
							07		24		00		17		25	
5 Group activity									—09	—42	—08	06	—22	49	33	—.19
									55	.11	30	.05	47	61	62	65
									63		14		—.27		40	
6 Combatant											70	17	18	—18	07	15
											43	.04	43	51	39	35
											24		—28		.07	
7 Offers invitation													.19	22	36	—51
													56	00	.53	—17
													02		01	
8 Accepts invitation															19	—.03
															.69	.55
															—48	
9 Constructive social																

considered sufficiently homogeneous to combine groups All analyses were conducted for each school separately.

In four schools, *B, H, I* and *K*, coefficients of rank difference correlations were computed pairing each of the categories of behavior with each other In this way a preliminary survey could be made of the consistency of found relationships In general the distributions of scores (and ranks) were greatly distorted for the various subdivisions of the behavior considered "violence," both aggressor-type and recipient-type, due to the infrequency of the observation of such behavior The intercorrelations that do appear as a result of pairing these sub-categories with each other are, in general, positive Thus, the pairing of Physical Violence as an "Aggressor" with Physical Violence as a "Passive Recipient" in the four schools were .51, 26, .66 and .42, all positive (see Table 1, categories 6 and 8). On the basis of such considerations, the further analysis was undertaken combining all the categories of the Violence rôle, both as aggressor and as recipient into a single score called "combatant" behavior. In five other schools *C, E, G, L* and *M*, the analysis of patterns was continued Here nine categories were considered (Violence Behavior being considered as one) and product-moment correlation coefficients were computed, again pairing each category with every other category of behavior (Table 2)

If we study these 36 pairings from the point of view of consistency among the nine schools (rather than the customary criterion for statistical reliability) we find only seven sets of correlations where the coefficients all have the same sign These pairings are

Sedentary and Attentive Interest in Others (correlates positively).
Sedentary and Offers Invitation (correlates negatively)
Sedentary and Constructive Social Use of Material (correlates negatively)
Attentive Interest in Others and Constructive Social Use of Material (correlates negatively).
Combatant and Offers Invitation (correlates positively)
Combatant and Constructive Social Use of Material (correlates positively)
Offers Invitation and Accepts Invitation (correlates positively)

The criterion of consistency of the sign of coefficients is used rather than reliability of the individual correlations because of the customary organization of preschool groups into numerically small units. When dealing with small groups, correlations are greatly influenced by one or two cases For the sake of an interpretative discussion of personality, the actual size of a correlation is of importance and its reliability deserves demonstration. For the present we do not feel that the conditions under which the data were collected were such as to warrant combining all the schools and deriving a single coefficient. Rather than test each coefficient and combine probabilities, the arbitrary criterion of agreement in sign was selected. This criterion may prove more interesting to nursery school education where the small units are maintained and within the framework of which educational theories and practices are developed. Interpretation of analyses of a matrix of intercorrelations by any of the various mathematical devices being still uncertain and not universally accepted, the simpler analysis here undertaken seems justified.

These seven pairings which seem to be consistently intercorrelated provide the basis for first considerations of behavior organization in preschool children. Further study of the pairs chosen from the five schools of Table 2 appeared justified by inspection of the same pairings in the four schools of Table 1 Because of the heterogeneity of the samples studied with respect to chronological age and the

TABLE 3

PARTIAL CORRELATION COEFFICIENTS BETWEEN SELECTED CATEGORIES OF BEHAVIOR (CHRONOLOGICAL AGE PARTIALED OUT)

	C	*E*	*G*	*L*	*M*
Sedentary and interest in others	12	02	67	14	49
Sedentary and offers invitation	—20	—.24	—52	—08	—36
Sedentary and constructive use of material	—33	—50	—22	—32	—38
Interest in others and Constructive use of material	—44	—05	—20	—50	09
Combatant and offers invitation	70	29	.62	48	.04
Combatant and constructive social use of materials	.02	23	.43	25	.00
Offers invitation and accepts invitation	.20	44	.65	—02	26

tremendous influence of such variation, these relationships cannot be accepted without further analysis. Partial correlations were resorted to for this purpose and Table 3 presents the intercorrelation of the seven pairings, for each of the five schools, with chronological age partialed out.

Again resorting to the criterion of consistency of sign of the intercorrelations we find that of the seven pairings, two pairs may be dropped from further consideration. These two pairs have a consistent intercorrelation through four of the schools, but in a fifth the relationship seems to be reversed when chronological age is partialed out. If two categories of behavior are not so related that when studied in school after school the correlations are in the same direction (either positive or negative), it does not seem as if students of child behavior can place much emphasis on the categories. This does not imply that the category is unimportant nor that it is not worthy of study. It does mean, however, that for the present it does not appear in a constellation of behavior with sufficient consistency for us to gain insight into the rôle the particular behavior plays in the organization of the child's personality. Lacking such insight and having no information concerning its rôle in the genetic development of the child, detailed consideration at present seems unjustified.

From Table 3 we find that there are five pairings of behavior categories which have consistent relationship through the schools analyzed and which maintain their relationship even when the factor of chronological age is taken into account by partial correlations:

1. Sedentary behavior is positively correlated with Attentive Interest in Others.
2. Sedentary behavior is negatively correlated with Offers Invitation.
3. Sedentary behavior is negatively correlated with Constructive Social Use of Material.
4. Combatant behavior is positively correlated with Offers Invitation.
5. Combatant behavior is positively correlated with Constructive Social Use of Material.

The attitude toward the results must still be somewhat tentative from the point of view of interpretation in the light of personality theory. There are two sets of data which indicate that the adult social standards and logical arrangements cannot be applied to the organization of child behavior First, there is the general positive correlation between various aspects of combatant behavior, being an aggressor is positively related to being the recipient of aggression In view of the manner in which the data were collected, if a child were *in a fight at one moment he might be displaying aggressor* behavior and a half minute later, recipient behavior But the apparent inconsistency of positive correlation between violent physical aggression and being the passive recipient of aggression seems to need further study

The second point of interest is the positive relationship between behavior in the "violence rôle" (Combatant behavior), and such social activity as "Offers Invitation" and "Constructive Social Use of Material" From the adult point of view and ethical norms, these behaviors represent different categories and to find them positively correlated seems an inconsistency in behavior. Further study is definitely indicated to fix satisfactorily the picture of genetic development from the childhood pattern to the adult norm, meanwhile, child behavior, may best be conceived in the light of the organ*ization of the child personality.* Data such as *are here reported may be* interpreted as supporting reports of positive correlation between such behavior as sympathetic forms and aggressive forms (L. B. Murphy), in pointing to organization of child behavior in terms of "general activity" and absence of such activity. Being sympathetic is part of the pattern of "much behaving" as is being aggressive In the present study, "Sedentary" behavior and "Interest in Others" are positively correlated and are part of the pattern of little activity, "Offers Invitation," "Constructive Social Use of Material" and "Combatant behavior" are all positively correlated as part of the pattern of much activity Further, these two patterns of behavior are negatively correlated.

B

The analysis of the behavior presented above suggests that behavior of children may be conceived of as organized in two spheres (*a*) an active, outgoing mode which includes all forms of behavior requiring the children to make advances, and (*b*) an inactive mode, a refrain-ing from modes of behavior making advances, and an occupation with activities or forms of behavior requiring the making of few if any such advances. These may be truly discrete categories or merely ends of a continuum We found that five categories of the behavior studied were intercorrelated with sufficient consistency to permit further consideration of them as stable elements of a pattern Two of these forms represented the inactive mode, three the active Now we may turn to a consideration of the effect of nursery schools on this behavior Correlations were computed for each of the 13 schools, between each of the five categories of behavior and the length of time elapsed from the original enrollment of the child in the nursery schools and the date of the tenth observation made during the course of the study These coefficients are presented in Table 4

TABLE 4

Correlation between Selected Categories of Behavior and Length of Time in Nursery School

	Sedentary	Interest in others	Combatant	Offers invitation	Constructive use of material
A	—35	—51	46	30	21
B	—46	—27	—03	—05	21
C	—.28	—19	51	74	42
D	00	18	—28	52	51
E	—26	—28	10	29	.43
F	—09	—24	27	—40	—15
G	—27	—32	20	.47	11
H	—29	—54	—.08	22	44
I	—.37	—38	26	.21	29
J	—56	—26	.43	34	70
K	—.22	—47	06	46	.29
L	—10	—52	—23	30	—08
M	03	22	—22	—48	19

In general, Table 4 indicates that nursery schools seem to be decreasing the tendency of children to manifest behavior of the inactive mode and increasing the display of active forms. Thus, sedentary behavior is negatively correlated with length of time in the nursery school in 11 of the 13 schools, as is Interest in Others (the two representatives of the "inactive" mode). The Combatant category is positively correlated with nursery experience in 8 schools, Offers Invitation is positively correlated in 10, and Constructive Social Use of Materials in 11. Undertaking a more careful analysis of the effect of nursery school training by use of the partial correlation technique, to take account of the variation due to differences in chronological age of the children, similar results are found (see Table 5).

TABLE 5
PARTIAL CORRELATION BETWEEN SELECTED CATEGORIES OF BEHAVIOR AND LENGTH OF TIME IN NURSERY SCHOOL (CHRONOLOGICAL AGE PARTIALED OUT)

	Sedentary	Interest in others	Combatant	Offers invitation	Constructive use of material
A	—38	—51	.46	.31	21
B	—19	—15	—04	—14	12
C	—32	—23	50	72	34
D	09	23	—27	.46	.52
E	—30	—.25	17	—15	18
F	—.10	—19	21	—49	—08
G	—24	—24	20	39	03
H	—04	—27	15	—06	40
I	07	—21	10	06	12
J	—31	—85	35	36	80
K	—16	—46	07	05	25
L	18	—22	—.03	08	—65
M	.11	28	—22	—05	.14

This analysis confirms the suggestion in Table 4, for here we find "Sedentary" behavior correlated negatively with length of time in the nursery school (partialling out chronological age) in 9 of the 13 schools, and "Interest in Others" negatively in 11 schools "Combatant" behavior correlates positively with nursery school training in nine schools; "Offers Invitation," positively in eight schools; and "Constructive Social Use of Materials," positively in eleven schools. The effect of nursery school training, in general, seems to be to diminish display of "inactive" forms of behavior and to increase "active" manifestations

It is far from clear, at present, the precise interpretation which is to be placed upon the two categories here described. Do they represent a basic dichotomy comparable to the types described by Jung? Do they represent merely two phases of a unitary aspect of personality? Do they (or does this) represent a basic pattern which pervades the whole personality of the child, or do these findings represent merely superficial and transitory aspects of children's behavior during a small portion of the day and under the special conditions of group stimulation? What is the "value" of increasing manifestations of one form and decreasing the other?

Much verbiage can be supplied at the present time but the questions here raised and their numerous ramifications (including all problems of social development conceived perhaps as individuation of these general patterns into the complexly structured whole) are primarily problems to be answered by empirical study, and until such study can be undertaken it seems that further discussion might well be delayed

C

The present study may best be summarized, perhaps, not by an itemization of findings but, rather, by an emphasis on a few of the major implications for general understanding of child behavior.

Specific acts or forms of overt behavior on the part of young children are not to be interpreted as representing in the child a facet corresponding to the interpretation one might place upon such behavior if manifested by an adult in our general culture under "normal" social conditions.

Empirical evidence suggests the organization of young children's behavior is in terms of general activity-inactivity rather than patterns in accordance with social-ethical norms No data are available for present consideration of the questions of typology nor the all-pervasiveness of such observations to other aspects of children's personality.

The effect of nursery school training seems to be to increase active forms of behaving and decrease inactive forms. Until problems related to further consideration of the interpretation of these patterns are cleared up, an evaluation of this training in the light of personality as a whole seems premature (e g., are there inherited bases for a true dichotomy? If there are, what is the ultimate effect in terms of the integrity of the personality of making an inactive child active? What is the ultimate effect in terms of social adjustment and usefulness?)

D SUPPLEMENTARY NOTE

It may be worth noting at this point that the emphasis placed on the results of Tables 1 and 2 is not without some reservation. The discrepancies between Tables 4 and 5 suggest that had partial correlation coefficients been resorted to, the specific pairings selected as appearing correlated consistently might easily have differed. Chance

TABLE 6

COEFFICIENTS OF CORRELATIONS OF CATEGORIES OF BEHAVIOR IN THREE MORE NURSERY SCHOOLS (COMPARABLE TO TABLE 2)

	Moving around	Interest in others	Parallel activity	Group activity	Combatant	Offers invitation	Accepts invitation	Constructive use of materials
Sedentary	—52 .44 00	.25 73 16	.18 —00 06	.75 —86 —50	—30 —01 —10	—25 —27 26	05 00 —31	—56 —10 —15
Moving around		18 69 34	03 —17 —15	—46 —31 —94	—60 —04 .62	.24 13 08	—.20 00 .21	—24 —15 14
Interest in others			.50 .54 —52	—.23 —72 —14	—40 —16 —.47	—21 —08 .00	—25 .00 —01	—59 —16 20
Parallel activity				—10 —30 —18	—48 —02 36	—47 11 10	—16 00 01	—33 —17 —08
Group activity					42 .02 —.10	—15 51 —20	20 .00 45	.18 22 .20
Combatant						—13 47 —02	34 .00 —09	.36 .29 —30
Offers invitation							—11 00 —04	68 56 12
Accepts invitation								08 00 02

factors influencing the age of the entry of children might well have caused a change in the sign of the obtained coefficients.

This difficulty, it is presumed, would not seriously alter the findings as presented due to the emphasis placed upon consistency through nine schools Recognizing the limitations of relying on but nine schools, the same analysis (not the proposed one) was extended to three more schools, *A*, *D*, and *J*.

This further analysis raises the total number of schools considered to 12. The intercorrelations of the behavior displayed appear in Table 6 and are directly comparable with the material presented in Table 2

Table 2 yielded seven sets of consistently correlated pairs (cf. p 343) ·

1 Sedentary and Interest in Others (positive)

2. Sedentary and Offers Invitation (negative)

3 Sedentary and Constructive Social Use of Materials (negative)

4 Interest in Others and Constructive Social Use of Materials (negative)

5. Combatant and Offers Invitation (positive)

6 Combatant and Constructive Social Use of Materials (positive).

7. Offers Invitation and Accepts Invitation (positive).

Comparing these pairs in Table 6 may best be done in columnar form (Table 7).

TABLE 7

Pair	On Table 2	On Table 6
1	5+	3+
2	5—	2—, 1+
3	5—	3—
4	5—	2—, 1+
5	5+	2—, 1+
6	5+	2+, 1—
7	5+	2—, 1 zero

The criterion of perfect consistency would no longer select five of these seven pairings if we applied it to twelve schools as a whole. It is to be noted, however, that the trend is the same except in the case of Pairings 5 and 7.

The precise cause for these changes may be sought for in the distributions of the obtained frequency of recorded observations, e.g , in Pair 7, "Offers Invitation" and "Accepts Invitation," we find Table 6 indicating that they correlate —04 in School *D* and —.11 in School *J*. Further study of this reversal of the positive correlation shows that the average number of times behavior in the "Accepts Invitation" was recorded for each child in School *D* was only 2 and for School *J*, 0.0. Pair 5 also involves the "Accepts Invitation."

Further inspection of Table 6 lends support to the general classification suggested of active and inactive patterns. This further analysis of three more schools confirms the hesitancy expressed in the preceding section to placing an emphasis on the specific details of the particular pairings found consistently intercorrelated, and also lends confirmation to the general hypotheses there presented.

REFERENCE

1. MURPHY, L. B. Social Behavior and Child Personality: An Explanatory Study of some Roots of Sympathy. New York: Columbia Univ. Press, 1937. Pp. 344.

Bureau of Child Development and Parent Education
University of the State of New York
State Education Department
Albany, New York

The Journal of Genetic Psychology, 1939, 64, 353-358

SEASON OF BIRTH AND INTELLIGENCE*

Teachers College, Columbia University

R. PINTNER AND G FORLANO

During the last decade several articles appeared on this topic. Most of them have dealt with data gathered in the United States, some in Europe, all of them with results obtained in the Northern Hemisphere.

The present study deals with data obtained in the Southern Hemisphere and also with certain samples of children of low *IQ* obtained in the United States. We were interested in raising the question as to whether the same tendency found in the Northern Hemisphere, namely the slightly lower average *IQ* of children born in the winter would be found also in the winter months in the Southern hemisphere where the seasons are reversed. We were also interested in enquiring whether the lower *IQ* in winter would be found among the subnormal group alone, and furthermore whether the number of subnormal children born varied in accordance with the seasons

We are indebted to many individuals for their kindness in sending us their results of tests and in allowing us to use them for this study Dr. K. S Cunningham of the Australian Council for Educational Research sent us 938 cases from Victoria and Queensland, Australia. Dr Cunningham was also instrumental in obtaining for us many results from Mr H. T Parker of Tasmania and Dr Phillips of New South Wales. Dr Collman sent us the data for 1,969 cases from Victoria, Australia. Mr M J. Kok contributed 158 cases from Transvaal, Union of South Africa With regard to the low *IQ* cases in the United States, we wish to thank Mr. Charles Rosenthal for tabulating 8,985 *IQ*s very kindly contributed by the Department of Ungraded Classes of the New York City Public Schools, and we are also indebted to Mr Charles Rosenthal for copying the *IQ*s of 658 cases which the Training School at Vineland allowed us to use

DATA FROM THE SOUTHERN HEMISPHERE

The samples of data are not very large for this type of study, in

---

*Received in the Editorial Office on June 24, 1938

which only a very small difference in *IQ* is to be expected, as we have found from our studies in the United States. Furthermore, *they differ widely in manner of selection.* Two samples, *A* and *B*, seemed to be random samples of a more or less unselected population. The other samples were obviously groups of rather low *IQ* cases, in all probability children who were tested because they were not doing well in school or because they were suspected of being mentally deficient. We have grouped these altogether in one distribution.

For each of these samples we have calculated the mean *IQ* and *the standard deviation for each month.* In sample *A* ($n = 1{,}969$) the number of cases per month varies from 139 to 190, in Sample *B* ($n = 938$) the number per month varies from 47 to 107. In neither instance are the number of cases per month sufficient to establish *statistically reliable differences for the monthly means.* Sample B is rather small, and wide variation in number of cases per month as well as the wide variation in the standard deviation of the monthly means (from 13.80 to 17.50) make us very dubious as to the randomness of this sample. For these reasons we are not reproducing here our table of monthly means for these two samples.

Our next step is to calculate the means and sigmas for the four seasons. We have arbitrarily divided the seasons as follows: *Spring*: October, November, December, *Summer*: January, February, March, *Autumn*: April, May, June, *Winter*: July, August, September. We have called Autumn and Winter the *Cold Season* and Spring and Summer the *Warm Season*. Table I shows the means and sigmas for the four seasons and the two half-yearly intervals for Samples *A* and *B* separately and also for the combination of Sample *A* plus Sample *B*. Sample *A* is our largest sample, namely 1,969, and probably our best sample. It shows the lowest seasonal

TABLE 1

	Sample *A*		Sample *B*		Samples *A* and *B*		
Season	Mean *IQ*	Sigma	Mean *IQ*	Sigma	Mean *IQ*	Sigma	*n*
Summer	105.75	12.60	95.50	15.10	102.40	14.25	668
Autumn	104.70	13.15	95.60	15.05	101.55	14.50	757
Winter	104.65	11.75	95.80	15.85	101.55	13.85	778
Spring	105.90	12.20	95.95	16.50	102.80	14.20	704
Warm season	*105.60*	*12.40*	*95.70*	*15.80*	*102.60*	*14.25*	*1,372*
Cold season	104.65	12.50	95.90	15.10	101.65	14.30	1,535

mean in winter, thus supporting our previous findings in the United States. Sample *B* is a smaller sample and probably not so good. The winter mean *IQ* is not the smallest, and it thus does not support our previous findings. The differences between the seasonal means for Sample *B* are all very small. For the combination of all cases we note that the lowest mean *IQ* is for winter and autumn, both being identical. Combining the seasons into half-yearly intervals we find the cold season has a mean *IQ* of about one point lower than the warm season. The ratio of this difference to its standard error is 1.53. We have computed critical ratios for all differences between seasons and between half-yearly intervals. For Sample *B* none of these ratios approach 1.00, for Sample *A* four out of seven lie between 1.00 and 2.00, for the combined group five out of the seven lie between 1.00 and 2.00. None of these ratios satisfy the criterion of statistical reliability. All that we can say, therefore, about these data from the Southern Hemisphere is that they support somewhat our findings in the United States, namely that the lowest mean *IQ*s are found in the cold months and probably mainly in Winter.

Our remaining data from the Southern Hemisphere consisted of four samples all having rather low mean *IQ*s. They are very unsatisfactory and are really unsuitable for the kind of study we are pursuing. We know nothing as to how they have been selected. We did, however, combine all these cases and calculate monthly and seasonal means and sigmas and ratios. We give them here in Table 2 because they do not support our previous findings, although the fact that they do not seems to us to have little weight in view of the unsatisfactory nature of the sample. Only two out of seven critical ratios are above 1.00, and none approach 2.00.

TABLE 2

	Mean *IQ*	Sigma	*n*	Below *IQ* 60 *n*	Below *IQ* 50 *n*
Summer	69.40	14.20	546	133	35
Autumn	69.05	13.90	535	125	41
Winter	70.05	15.70	586	138	55
Spring	69.95	15.20	547	127	50
Warm Season	69.68	14.73	1,093	260	85
Cold season	69.57	14.87	1,121	263	96

One thing worthy of note about this sample of children of low *IQ* is that the largest number of cases, namely 586, were born during the winter months. We have also counted the number of cases below 60 *IQ*. These are shown in the next to the last column. Again we note the larger number of cases born in Winter. If we count those undoubtedly feebleminded, having *IQ*s below 50, we have the results in the last column. Here by far the largest number of cases are born in winter.

Low *IQ* Groups in the U. S. A.

The data here consist of 8,985 "special class" cases with a mean *IQ* of 65.20 and a sigma of 7.14, and a small group of 658 institutional cases with a mean *IQ* of 52.97 and a sigma of 19.31.

Table 3 gives the essential results for the main sample of 8,985 cases. The standard deviation for the total group is 7.14, indicating the restricted range of the sample. The sigmas for each month vary very little, ranging from 6.50 to 7.56. The monthly means range from 64.79 in February to 65.53 in March. The difference here is only 0.74 and the standard ratio is 1.94. With reference to the

TABLE 3

	Total cases			Male			Female		
Month	Mean	σ	n	Mean	σ	n	Mean	σ	n
January	64.83	7.56	836	65.15	7.60	547	64.25	7.60	289
February	64.79	7.30	742	65.07	7.20	460	64.35	7.70	282
March	65.23	7.25	801	65.48	7.40	508	64.80	6.98	293
April	65.40	7.40	695	66.25	7.60	394	64.30	6.98	301
May	65.53	7.04	716	65.80	7.60	444	65.10	6.00	272
June	65.36	6.50	680	65.70	6.70	420	64.82	6.00	260
July	65.23	7.24	759	65.71	7.25	517	64.23	7.20	242
August	65.34	6.94	787	65.54	7.00	509	65.00	6.95	278
September	65.20	7.52	797	65.33	7.50	491	65.01	7.50	306
October	65.26	6.97	782	65.73	6.60	474	64.56	7.50	308
November	64.89	7.53	674	65.62	7.45	398	63.84	7.60	276
December	65.46	7.04	716	65.42	7.10	425	65.52	6.95	291
General	65.20	7.14	8,985	65.55	7.27	5,587	64.65	7.13	3,398
Spring	65.46	7.00	2,091	65.91	7.33	1,258	64.72	6.38	833
Summer	65.23	7.24	2,343	65.53	7.26	1,517	64.78	7.24	826
Autumn	65.17	6.90	2,172	65.59	7.05	1,297	64.65	7.38	875
Winter	64.96	7.38	2,379	65.24	7.42	1,515	64.47	7.43	864
General	65.20	7.14	8,985	65.55	7.27	5,587	64.65	7.13	3,398
Warm months	65.33	7.13	4,434	65.70	7.29	2,775	64.75	6.82	1,659
Cold months	65.06	7.16	4,551	65.40	7.07	2,812	64.56	7.41	1,739

seasons we note that the lowest mean, 64.96 is for Winter and the highest for Spring, 65.46 The difference here is 0.50, and the standard ratio is 2.32, which is statistically significant. Winter has the lowest mean *IQ* for each of the sexes taken separately. The difference between the warm (April to September) and cool (October to March) half-yearly intervals is very small, but the mean for the cold season is lowest, and this is also true for the sexes taken separately.

With reference to the actual number of children born during each season, we find that the largest number is born during the three winter months of January, February and March, although these three months have the fewest number of days of all the four seasons

This large sample of children of low *IQ* therefore shows the same tendency as found in previous samples of individuals, namely a lower *IQ* in Winter

The other sample of 658 institutional cases is too small to warrant a detailed publication of monthly means We give in Table 4 the results for the seasonal groupings Again the differences are small, but the lowest mean falls in Winter, and the cold season has a lower mean than the warm season However, the largest number of cases is not found in the winter season

TABLE 4

	Mean *IQ*	Sigma	*n*
Spring	53.24	18.27	169
Summer	52.86	19.77	155
Autumn	53.22	18.93	174
Winter	52.22	20.62	160
Warm season	53.06	19.00	324
Cold season	52.74	19.36	334
Total	52.97	19.31	658

SUMMARY

Data from the Southern Hemisphere tend slightly to confirm previous findings in the Northern Hemisphere, namely that the mean *IQ* of children born in Winter is somewhat lower than that of children born in the other seasons The samples of cases we have been able to procure are not large and in other ways are not very satisfactory. It is to be hoped that larger and better samples may be studied in the future.

The data for the low *IQ* groups in this country also show a lower *IQ* in the winter season.

Hypotheses as to the reason for these trends have been suggested in previous articles by the present writers as well as by others. They will not be repeated here. So far no eminently satisfactory hypothesis has emerged.

Teachers College
Columbia University
New York City

The Journal of Genetic Psychology, 1939, **54**, 359-390.

AN INTELLIGENT-QUESTION METHOD OF TEACHING AND TESTING*

Department of Psychology, University of Wisconsin

HULSEY CASON

A INTRODUCTION

The general purpose of the present method of teaching and testing is to enable the student to learn as much as possible with a minimum of wasted time and effort The method may be used in connection with the lecture-reading-discussion part of many university, college, and high school courses, but it will probably be found of most value in connection with the subject matter of cultural courses. The influence of academic courses is sometimes principally a function of the scholarship and personal qualities of the instructor, but assuming that these factors of personality are equalized, our procedure seems to represent some improvement in the methods of both teaching and testing.

The intelligent-question (*"I-Q"*) method is largely concerned with the teaching of meaningful material, and our main thesis is that the psychological processes of the students which are tested by the examinations should be the same as those which the instructor deliberately attempts to cultivate in the students by his teaching We are principally interested in the internal-organic (psychological) processes which are present in the students while they are listening to lectures, reading, taking part in class discussions, preparing for examinations, and taking examinations The present *I-Q* method will have a special value for those students who are not entirely clear on what material they are expected to learn in a course, how they are supposed to learn it, the relative emphasis they should place on different aspects of the subject, and the quality of organization and mastery of detail that will be required on the examinations.

B GENERAL PSYCHOLOGICAL PRINCIPLES

After using the present method for several years, and after attempting to describe it in a systematic manner before several academic

---

*Received in the Editorial Office on June 27, 1938

groups, we have become aware of several factors which make a straightforward presentation of the problem quite difficult. The majority of university teachers are not particularly interested in the literature on the methods of teaching and examining, and they commonly assume that this is a topic on which they may have free opinions in somewhat the same way that they have beliefs which are pleasing to them on politics, religion, and morals. Some of the psychologists and educationists who are interested in tests and who may have tests to sell have strong attitudes on the value of the tests, and they may be sensitive in regard to the apparent weaknesses of the tests. Finally, some of the leaders of the testing movement have lost interest in several of the more or less elementary psychological principles on which the theory of teaching and testing depends.

We shall take up the general psychological principles on which the theory of teaching and testing is based in order to avoid the fairly common loss of intellectual bearings. The psychological principles described in the following paragraphs are the basis of our *I-Q* method (Sections *C* and *D*) and our later discussion of "Test or Tact?" (Section *E*). It is necessary to have these psychological principles *in mind before the I-Q method of teaching and testing can be clearly* understood, and we would like to ask the reader to decide whether *he does not accept each of these principles, and also whether our* method and the advantages of our method do not follow naturally from these principles.

1. *The Absence of Psychological "Elements"*

In one of the historical phases of psychology, the mentalists assumed that all conscious psychological processes could be reduced to elementary processes, such as sensations, feelings, and images, and the Russian reflexologists and American behaviorists assumed that all forms of behavior were made up of simple reflexes. These assumptions of atomistic psychology have now been discredited and discarded, and few if any pure mentalists and reflexologists remain.

The first intelligence tests were devised at a time when atomism was the current psychological doctrine, and the modern writers who speak as if intelligent meaningful material can be broken down into its elements include many psychologists and educationists who are or have been engaged in the task of constructing psychological tests.

There are, however, no clear elements to which much of the intelligent material learned by college students can be reduced

A sentence is not equal to the sum of the disarranged words in it, and a paragraph is not the same as the jumbled sentences Poems, stories, essays, and plays are not made up of so many little details compounded together In a sentence there are psychological differences between nouns and verbs, paragraphs have a certain amount of unity, and stories have a plot Most of the material that we read, hear, think, write, and speak about has a certain amount of intelligent organization.

Instead of attempting to reduce the subject matter of academic courses to elements which do not exist, it is generally necessary for one intelligent person to judge the intellectual product of another individual, and it seems that the greater a person's mastery of the important facts and principles of a subject, the better he is able to judge what another person writes or says about the subject. It is sometimes as difficult to analyze one's intelligent judgment process as it is to analyze a literary composition or a musical selection Many important intellectual subjects are being taught in the schools which have to be judged but which cannot be broken up into disconnected items. The conscious judgment process is itself somewhat elusive from a scientific standpoint; and attempting to analyze one's own judgment process may injure the process in somewhat the same way that a skilled manual performance may be disturbed by paying attention to details of which one is ordinarily not clearly conscious We are quite far from an adequate understanding of the judgment processes of students in taking psychological tests, or the choice processes of psychologists in selecting the items for the psychological tests

2. *Learning, Retention, and Reproduction of Meaningful Material*

All of the important and interesting topics in every field of knowledge have a number of meaningful relations, and these relations give to the facts and principles whatever importance and significance they may have The most efficient method of learning organized material is to think about it and observe the relationships in it, and knowledge of a topic consists principally in the understanding of these meaningful relationships The instructor generally has a good memory for the more important and interesting facts of a course, and these are the topics about which students most frequently ask questions The more meaningful facts and principles are the material which both the instructor and the students generally wish to learn and remember (Whipple, 1915, 205-211; English, 1934; Welborn and English, 1937).

The facts and principles of a course should be taught in as meaningful a setting as possible, and no type of examination should be given which encourages the students to learn detailed facts out of or in any ways separate from their proper intellectual setting

3 *The Psychological Center*

The intellectual processes of observing, understanding spoken and written language, thinking, speaking, and writing make up what may be regarded as the psychological center. People generally talk to themselves while they read or listen, think while they speak or write, hear what they themseves say, read what they themselves write, and so on An individual may learn by reading or by listening to another person speak; and if he has a good understanding of what he has learned, he can think about the material and can express himself by writing or by speaking about it Most of what people learn and remember is learned and remembered in connection with these conscious-verbal processes of the psychological center.

The principal aims of most university courses are to enable the student to put the subject matter of the course in his psychological center in such a way that he can (*a*) understand, appreciate, and judge what he reads about it and what he hears other people say about it, (*b*) think in an integrated way about it, and (*c*) speak and write intelligently about it The general purpose of the method of teaching and testing which we shall describe is to arrange academic conditions in such a way that the student will put the subject matter of a course in his psychological center and will put it there in the most efficient and sensible manner

4. *The Form-of-Learning Rule*

There is a valuable general rule which states that everything should be learned in the form in which the individual will later desire to use it. The pattern learning situation is often important, and the reproduction of learned material may be seriously disturbed by the introduction of unaccustomed distracting stimuli or condi-

tions or by the withdrawal of habitual associated stimuli or conditions The reproduction score may be markedly influenced by the kind of reproduction test that is employed. The correlation between different kinds of learning is quite low, and one form of learning frequently transfers only to a very limited extent to other *forms of learning*

The psychological processes for which the students are tested on the examinations should be the same kind of psychological processes as those which the instructor has deliberately and overtly attempted to produce in them by his teaching The students should know what *kind of psychological processes will be valued and credited on the* examinations, and they should be given every possible opportunity and aid in deliberately cultivating these psychological processes in their work for the course. In other words, those psychological processes that are taught should later be tested, and those that are tested *should have been taught*

The present form-of-learning rule does not seem to be clearly understood by many teachers, and it is sometimes more or less deliberately violated in the forlorn hope of cultivating some kind of "power," "ability to think," or originality in "using" material *acquired by the straightforward processes of learning*

It is very easy to find examples of what one should and should not do in this matter of teaching and testing One should not spend the time of a course in having students translate French into English, and then as a test for "originality," ask them to translate English into French on the final *examination; and it would be unwise* and unnecessary to have students recall the meaning of German words in context during the term, and then as a test for "ability to think," ask them to recognize the words out of context on the examinations In a history course, there would be no point in spending all of the time on the *military history of a country, and* then a test for "power," ask only about social customs and educational institutions (Brinkley, 1924, Hawkes, 1936, 204-209), and in an economics course, one would not be justified in spending the semester learning generalities and no details, and then in the hope of determining whether *the students can "use" what they have* learned, ask about details rather than generalities (Kelley and Krey, 1934) If disconnected details are tested on the examinations, disconnected details should have been taught during the term, and if original thinking is tested on the examinations, original thinking should have been taught in the course If intelligent material is taught in the course, the examinations should be concerned with intelligent material Students have a right to know what they are supposed to learn, and there is no reason for taking them by surprise on any examination.

The present rule that one should teach what will be tested and test what has been taught is opposed to the use of most of the traditional new-type (*N-T*) tests (Sims, 1932). The *N-T* tests seldom test what has been taught, and they have very little value as teaching devices (Hamilton, 1929, 86-112) If an instructor wished to teach the boiling point of water by the multiple-choice method, for example, I presume he would have to tell the class that the boiling point of water was 192°, 202°, 212°, 222°, or 232° *F*, have everybody guess at the boiling point of water, make up a distribution table of all the guesses, and then correct for the factors of suggestion and chance If an instructor uses *N-T* tests as substitutes for examinations, he should also teach by the *N-T* method, but if intelligent material is taught in a course, intelligent questions should be asked on the examinations.

5. *Significant Differences between Recall and Recognition*

Recall and recognition are the two principal forms of the conscious reproduction of learned material, but of all the different forms of memory, recognition is the simplest and most elementary. In a familiar experiment on recognition, 20 words are successively shown the subjects, and each word is exposed for a period of five seconds during which time the subjects think about and try to form some associations with the word. The just seen words are then mixed with 20 other words which have not just been seen, and the subjects are asked to state whether or not each of the 40 words has just been shown In a test of this kind, it is not uncommon for the majority of a group of subjects to make practically a perfect score

People can recognize many things which they cannot recall, but if an individual can recall anything he can practically always recognize it Recognition is especially superior to recall in the case of material that is rich in associations.

Several of those who are engaged in test construction believe that there are no significant differences between the psychological processes of recall and recognition, but these differences are quite numerous and varied, and many of these differences are of basic importance for the theory of tests and examinations Various differences between recall and recognition have been reliably demonstrated with several different kinds of subjects, with various types of memory material, and in a great variety of experimental situations, by Wolfe (1886), Lehmann (1889), Hoftding (1893), Gamble and Calkins (1903), Katzaroft (1911) Claparède (1911), Hollingworth (1913), Heine (1914), Owen (1915), Woods (1915), Strong and Strong (1916), Wheat (1923, 64-86), Symonds (1926), Gates (1927, 116-140), Rabaud (1928), MacCurdy (1928, 112-130), Seward (1928), Woodworth (1929, 105-109), Wylie (1930), Cason (1932, 90-91), Jenness (1932), Pillsbury (1934, 506-516), Clarke (1934), and Hanawalt (1937) These studies have shown that recall and recognition are not the same psychological processes and that recall and recognition tests do not measure the same psychological functions

The correlations are quite low and sometimes negative between different forms of learning (Hall, 1936, Husband, 1939), and most of the correlations between recall and recognition (even when the tests are based on the same memory material) are also quite poor (Myers, 1914, Achilles, 1920; Lee, 1925; Henmon, 1929, 14-18, 346-353, Bolton, 1931, Anastasi, 1932) For our present purposes, these results mean that if the students' learning processes in a course are directed principally towards recall, the examinations should also be concerned with recall

Thinking involves the processes of both recognition and recall, but it is principally a process of recall, and recall is the more important process in the student's acquired ability to think, speak, and write intelligently about meaningful material. The Psychological activities involved in our *I-Q* method are largely the processes of recall, thinking, and understanding; and such processes as are involved in answering or guessing at the traditional *N-T* tests, especially the true-false, multiple-choice, best-answer, and matching tests, are largely processes of recognition.

Since recall is the principal process that students are urged to strive for in most university courses, recognition tests should not be used on the examinations; and the very poor correlations between the different forms of learning, and between recall and recognition, strongly indicate that one cannot assume without evidence that just any kind of recall or recognition test will correlate (even moderately well) with a university student's intelligent understanding of a subject

6 *The Validity of Tests*

The reliability of a test is the extent to which it measures the same thing on different occasions and therefore correlates with itself; the validity of a test is the fidelity with which it measures what it purports to measure (Long and Sandiford, 1935), and all good tests are supposed to possess a certain degree of both reliability and validity It is necessary to have a criterion in determining the *validity of a test, but the test itself is always a poorer measure than* the criterion Tests never have validity in general, but always validity for a particular purpose

The Dean of a law school, for example, might want students who will do well in his law school, and grades in law school may be used as the criterion in selecting test items for a test which it is hoped will later show some correlation with grades in law school In selecting items for such a test, a large number of psychological test items may be given the law students, and a selection may then be made among those test items which show some correlation with the criterion of grades, probably omitting some of the test items which correlate well with each other The test items which have been selected in this way may be regarded as a general test for proficiency in the law school, although such a test will probably not correlate well with the reasonable criterion of grades But even if the test items selected in this way did correlate fairly well with the criterion of grades and therefore have some validity, the selected test items themselves cannot be as good a measure of proficiency in law school as the accepted criterion of grades

After using grades as the basis of selecting the psychological test items, one cannot then separate himself from his logical bearings and assume that the psychological test items selected in this way are a better measure of proficiency in law school than the grades.

There has been an almost unbelievable amount of loose thinking on this all-important question of the validity of psychological and educational tests, and this basic question of validity has been ignored in most of the claims for the value of the traditional *N-T* tests.

The validity of a psychological test is measured by the raw correlation between the psychological test and some sensible criterion such as grades or teachers' estimates, and not by the boosted correlation between the psychological test and some other similarly constructed test involving similar psychological processes It is possible for psychological tests to correlate highly with each other and still have no validity or value Many psychological tests have been correlated with the Stanford Revision of the Binet-Simon Test, using the Stanford test as the "criterion," and assuming that in this way some light would be thrown on the basic question of validity, but this "criterion" has been poor at the upper and lower ranges, it has no clear meaning above the age of about 16, it has not itself been validated in the usual manner, and it has recently been revised (Terman and Merrill, 1937) *It is necessary to have a sensible criterion in mind and to keep it in mind in judging the validity and value of any test* (Skaggs, 1927; Piéron, 1928).

In a few, but very few, studies of the traditional *N-T* tests, some attempt has been made to correlate the scores on the *N-T* tests with the reasonable criterion of grades on intelligent examination questions or with teachers' estimates, and when the validity of the *N-T* test is found to be poor, which is practically always the case, it has been claimed that the *N-T* tests are very good, but that grades on intelligent examination questions are in this way shown to be unreliable But the *N-T* tests cannot be as good a measure of a student's knowledge of a subject as the criterion of the intelligent examination questions themselves

In the heat of the argument about the alleged unreliability of the intelligent examination questions, some of those who are interested in constructing tests have forgotten that the *N-T* tests themselves have a very low validity

7 *Are There Any Fixed Psychological "Abilities"?*

No living organism has any anatomical or physiological character which remains completely constant throughout the whole period of life, and the evidence against any fixed psychological "abilities" and in favor of a moderately fluid conception of human nature is quite convincing Some of the data supporting the moderately fluid conception of human nature are found in observations on the superiority in intellectual achievement today as compared with that of several centuries ago, in the studies of identical twins reared apart, in the investigations of the canal boat children, and in the effects of sensory deprivation Research in technical genetics and embryology has discredited the traditional notion of "potential abilities" which were said to fix the upper limits of "development."

If a student receives a low grade in a mathematics course, he should not conclude, and he should not be encouraged to believe, that he lacks the necessary "ability" to learn the mathematics. He should not be taught an unnecessarily discouraging psychology He should be encouraged to believe that he can markedly improve his mathematical performance (*a*) by attaining a better mastery of the basic mathematical operations, and (*b*) by working harder and spending more time on the course he is taking

The most intellectual-sounding arguments in favor of "fixed" psychological "abilities" are offered in connection with such abstruse matters as Spearman's "*G*" which few can understand and which in the end generally turn out to be non-psychological subjects The idea that all of the known and unknown traits of school children just "grow" like a plant in a pot has been overdone. Although it is not yet possible to change some of the small percentage of feebleminded in the population into normal people, it is important to remember in connection with one's educational philosophy that all kinds of cultural achievement can be radically influenced by favorable educational factors

C THE INTELLIGENT-QUESTION METHOD OF TEACHING AND TESTING

The general psychological principles which we have described in the preceding section constitute the basis of our intelligent-question (*I-Q*) method which is briefly outlined in the present section These general psychological principles are also the basis for several of the advantages of this method (Section D) and for our later discussion of "Test or Tact?" (Section E) The principal features of our method of teaching and testing are as follows

(1) The students are given a mimeographed list of the 10, 20, or 30 topics which will be covered in the course, and during the term their attention is frequently called to the relation of the particular topic being considered to the whole organization of the course

(2) The instructor makes out the kind of questions that he desires for each topic, and the answers to these questions include

everything the students are asked to learn from the lectures, reading, class discussion, and from any and all other sources The 100 or 150 questions on the subject matter of the whole course are numbered consecutively, and all of the questions on the course or on each topic are mimeographed and given to the students in advance.

The subject matter of every course is naturally limited, but a university instructor frequently decides on some of the topics and subtopics which he will include in his course, and he often has some liberty in regard to the relative amount of time which he will spend on each topic Giving the students a list of topics and a list of questions on each topic may be regarded merely as a convenient means of marking off the subject matter of a course

(3) Whenever the instructor is lecturing or giving a demonstration and whenever a class discussion is in progress, the students always know which topic and which question are being considered, and during the lecture the instructor should in one way or another always be presenting one of the topics and attempting to answer one of the questions It may be desirable for the instructor to draw out the students from time to time to be more certain that they understand the terms he is using and the topic being presented, and the students should be more or less at liberty to interrupt the lecture and have a discussion at any time.

(4). Whenever it is possible, one or more references may be given on each question, indicating where some material on the topic can be found, but in answering each question, the students are urged to use any pertinent information they may obtain from any source (Stalnaker and Stalnaker, 1934) All of the references on all of the topics may well include several times as much material as most of the students will be inclined to read. This procedure gives a certain amount of flexibility to a course, and permits students of different interests and various degrees of preparation to do much more work on some topics than on others

(5) In order to save time, some of the topics covered in the text and in collateral reading may not be taken up in the class meetings at all, but the students should have a list of the questions on the text and other references while they are reading for the course.

(6) It will probably be desirable for some of the 5, 10, or 15 questions on each topic to refer to other subtopics in the same topic, in a different topic, or in several topics.

(7) The relation of each subtopic to the topic and the relation of each topic to the whole organization of the course should be kept clearly in mind at all times

(8) The selection of all questions for the examinations during the term, for any make-up examinations, and for the final examination is limited to the list of consecutively numbered questions covering all of the topics and subtopics of the course

D ADVANTAGES OF THE INTELLIGENT-QUESTION METHOD OF TEACHING AND TESTING

Several features of our method of teaching and testing are based on the general psychological principles described in Section *B*, and the following seem to be the principal advantages of the method which has been briefly outlined in the preceding section.

(1) The students always know the names of the topics and subtopics which have been taken up, which are being taken up, and which will be taken up This knowledge has the same value for them that the titles of the sections and subsections of a book have for the reader of the book.

(2). The students become familiar with the general setting and meaning of a question at the time it is first taken up in lecture, if it is taken up in lecture; and if there is any doubt about the way the question is phrased or the adequacy of the instructor's treatment of the subject, these matters can be considered at the time the subject is first presented.

(3). When each subtopic is taken up, the students always know what they will be expected to learn, and whether the question which they may later be asked is a "what," "why," or "how" question, whether they may be asked to define, give or state, name or list, describe or discuss, compare or contrast, distinguish between, evaluate, summarize, generalize, give a detailed analysis of, support the statement that, or give arguments pro and con (Hamilton, 1929, 136-156) The students should know in a general way how the subtopic on which a question is based should be learned while the original learning is taking place, that is to say, while they are listening to a lecture on it, engaging in a discussion of it, or reading about it

(4) Our *I-Q* method avoids the discouragement which sometimes results when a student studies passably well for an examination and then finds that he has learned *the wrong things or has learned them* in the wrong way.

(5) When the student knows in a general way what he is supposed to learn and how he is supposed to learn it, he can take notes more intelligently during the lecture and while he is reading, and the notes which he takes will also have a much more assured value

(6) The commonly used expression "essay question" is not adequate to cover the various forms which our intelligent questions may take, several examples of which have been given in (3) above The verb "essay" means to try to do or accomplish, especially something difficult, and the noun "essay" means a literary composition on a special topic

The word "intelligent" is the most suitable qualifying adjective for the kind of questions we have in mind, and we have made a free use of the expression "intelligent-question" or "*I-Q* "

(7) In answering each question, the student is urged to use any pertinent information which he may have obtained from any source, to maintain a refined and appreciative, but critical, attitude, and to exercise his higher psychological processes in any reasonable way that he may desire

In courses where the lecture-reading-discussion method is employed, it frequently happens that there are several equally good but different answers to a question, and it is assumed that cut-and-dried rote answers will be the exception rather than the rule. There should be some difference in procedure when a very competent instructor is teaching a factual subject to beginning students and when some other instructor is teaching a speculative subject to advanced students

(8) In giving an examination with the present *I-Q* method, all the instructor has to do is to read off a selection of the numbers of the questions to be answered

(9) The students are generally not taken by surprise by any of the questions because they have seen and thought about all of them before, they write the answers with considerable readiness, and do not waste time wondering what any of the questions are supposed to mean

(10) Since the students have already seen all of the questions, none of the questions have the appearance of catch questions, and the influence of the factor of chance in selecting examination questions is greatly reduced.

(11) Since the students have already heard or read, and thought

about, the answers to all of the questions before they come to the examinations, they have a fairly good recognition memory for what may be reasonably regarded as the correct answers, and they know the approximate grades which their papers will receive

(12) Inasmuch as the factor of chance is greatly reduced and since practically all of the students know what they have been asked to do and what they are doing, the instructor, in grading a student's answer to a question, is in a better position to judge the student's acquired ability to present the material in an integrated fashion, his intelligent understanding of the subject, and his appreciation of the meaningful relationships that are involved

In grading a student's answer to a question, the reader of the paper has to have in mind some kind of absolute standard of the answers which he considers should correspond, for example, to grades of *"A,"* *"B,"* *"C,"* etc ; for the only alternative to this absolute-standard system of grading is to give blind letter grades on the basis of a meaningless distribution of figures which have no passing grade

(13) Instead of having the "normal law fetish" as an ideal, the general purpose of a course should be to have the students learn the answers to the questions so well that practically all of them will deserve and receive high grades If they are not proficient in answering the questions on one examination, they should be requested to learn more about the subject and raise their grade on a second examination It is generally a matter of no importance at all whether students learn a certain block of course material in the month of February, for example, rather than in the month of January.

If very few of the students know anything about the subject at the end of the term, it is very probable that the instructor is failing in his work Neither endowed colleges nor tax supported universities exist for the purpose of giving tests and finding out what the students do not know

(14). Our *I-Q* method of teaching and testing tends to discourage the lecturer from wandering from the subject he is supposed to be presenting and from the field in which he is qualified to have some kind of intelligent opinion, and the students are saved from some of the more oratorical lectures and appeals to the gallery

(15) Regardless of the methods of teaching and examining that are employed, students frequently ask the instructor what they are supposed to learn and how they are supposed to learn it; and

in our *I-Q* method the instructor can answer all of these questions with some care when he first writes out the questions on the subtopics The instructor can in general and in particular stress the features or parts of the course that he desires to stress

(16) Although it should be clearly understood that equally good answers to a question will often take a variety of forms, there is a practical advantage, especially in connection with some subjects, in having the instructor attempt to give a clear, intelligent, and complete answer to a question before he asks the students to answer the question.

(17) In the teaching of almost any course, it is necessary for the instructor to take some kind of position on the value of straightforward learning as compared with attempts at original thinking A good many of the arguments against "killing the student's initiative" and in favor of encouraging him "to think" are rationalizations on the part of instructors who do not have an excessive amount of factual material to teach, and rationalizations on the part of students who are not overly inspired to learn The facts and principles of many subjects are so well known that original thinking on the part of the students and the instructor would be a waste of time It is necessary to know something about a subject before one can think efficiently about it, and persons of very limited knowledge may not realize what their thinking about a subject resembles The nice-sounding concepts of "creative imagination" and "ability to think in general" exist principally in the form of concepts.

(18) In all that we have said about the *I-Q* method, it is understood that the instructor will urge the students to exercise their higher thought processes in any and all of their work for the course; but it is also assumed that the students' knowledge of the subject on which they are taking a course is limited, that they have a desire to learn, that what they are asked to read is worth reading, and that the instructor is in some measure competent to teach the course Finding answers to the instructor's carefully formulated questions and listening to his lectures should not "kill the student's initiative" any more than reading in the assigned references, and there is generally no special reason why the instructor or the author of the text should take an overly self-effacing attitude

The great majority of instructors have an all too common difficulty in trying to persuade their students to do any original thinking at all, and the practical problem is much more frequently concerned with what is the necessary precondition of original thinking, that is to say, the acquisition of a good store of accurate and well organized information about a subject.

E. Test or Tact?

The *N-T* short-answer tests seem to be used in less than 10 per cent of all of the courses in the Letters and Science colleges of the United States, and they have hardly been experimented with at all in European countries. They are much more common in normal schools and teachers colleges, and they are frequently employed for experimental purposes by those who are more interested in efficiency than in general culture (Hawkes, 1936, 381-440), but some use has been made of the traditional *N-T* tests practically all the way from the graduate school to the crib (Ruch, 1929, Cason, 1931). The true-false test is the most widely used and discussed of all the *N-T* tests, and it is also the most objectionable *N-T* test.

Some of those who have published manuals on the *N-T* tests (Paterson, 1927, Orleans and Sealy, 1928) have assumed that it was no longer necessary to give any arguments in favor of these tests. Books and manuals on the *N-T* tests are widely used at the present time, although since 1931 more articles have been written favoring intelligent examination questions than the *N-T* tests. With the exception of Odell's (1928) book which is very good in many ways, the best books on examinations in the library (for example, Monroe, 1931, 1936, Jones, 1933; Desclos, 1936) seem to be the books that are least frequently used.

1. *Have the New-Type Tests Any Value in the Business of Teaching?*

Our *I-Q* method of teaching and testing is quite different from the traditional *N-T* tests in both spirit and procedure, and if a measurement-conscious teacher believes the *N-T* tests have any value in the business of education, he will probably be somewhat critical of our method. On the other hand, if a person interested in teaching sees any advantages in our *I-Q* method, he will probably feel that the traditional *N-T* tests have relatively little value.

In contrast with those who speak enthusiastically about a "syncopated sampling" of *N-T* items and who seem to derive more pleasure from testing than from teaching, we shall in the following discussion attempt to show that the traditional *N-T* tests place a special premium on the art of guessing (Brinkmeier and Keys, 1930, Seibert, 1932, Hill, 1937), that they are not based on any sound psychological principles, that they are not "objective," that they have very poor validity, and that their influence on the students is undesirable. In other words, we believe that the traditional *N-T* tests have no value whatever in the business of education, that their use should not be permitted in the schools, and that no academic credit should be allowed when they are used exclusively in the place of examinations.

The custom in some quarters of teaching organized and meaningful material, and then, on the examinations, of not permitting the students to tell what they have learned (Kimball, 1928) but of requiring them to judge or guess at a large number of disconnected items, seems to be an educational fad which will pass out of use after a few more years.

2. *Are the New-Type Tests Based on Any Sound Psychological Principles?*

In the theory of the *N-T* tests, the assumption is generally made that it is possible to break down the important subject matter of a cultural course into its "elements." The *N-T* tests do not encourage the students to learn important facts and principles in such a way that they can think and speak and write in an integrated way about these facts and principles (Jersild, 1929, Chen, 1933, Bagley, 1937). The *N-T* tests do not encourage the intelligent and serious students to observe and think about the various meaningful relationships that are present in the subject matter of most university courses. They violate the form-of-learning rule that a person should learn everything in the form in which he will later desire to use it, and they at the same time violate the principle that one should teach what will be tested and test what has been taught. It would be a valuable contribution if someone would describe the sound psychological principles on which the *N-T* tests are based. I have looked for these principles, but have not been able to find them.

3. *Are the Traditional New-Type Tests "Objective"?*

The *N-T* tests are often spoken of as "objective" tests, and it is true that such recognition tests as the true-false, multiple-choice, best-answer, and matching tests may be scored without error if the items are not weighted for importance, but the completion (recall) test, which in some ways is the least objectionable of all the *N-T* tests, is quite difficult to score and cannot be scored without considerable error

The word "objective" is an antonym for "subjective" in the "subjective-objective" dualistic dichotomy, and during the recent wave of enthusiasm for behaviorism, the mere statement that anything was "objective" sometimes settled all of the arguments in its favor The items of a test have to be selected and put in the test, and if several different people make out *N-T* tests on the same intelligent material, the *N-T* tests of the different people tend to correlate quite poorly with each other (Pullias, 1937) Some of the supporters of the *N-T* tests seem to have forgotten that the items of the *N-T* tests have to be put in the *N-T* tests

When the teacher of a course selects particular items to put in an *N-T* test, he uses his own conscious-verbal choice processes in selecting the items, and these are not "objective" processes, and when students guess at or attempt to answer the items of an *N-T* test, they use their own conscious-verbal judgment processes, and these are also not "objective" processes It would be quite difficult for a teacher to select items for an *N-T* test without the use of his conscious-verbal choice processes, and it would theoretically be difficult for students to take an *N-T* test without the use of their conscious-verbal judgment processes The construction of all examinations and the taking of all examinations are "subjective" processes since the conscious-verbal processes of the teacher and the students are always involved

4 *Have the Traditional New-Type Tests Any Validity?*

The one objection to the *N-T* tests that is sufficient in itself is that they are not valid measures of what the administrative officers and faculties of universities want the students to learn (Cason, 1931; Kelley and Krey, 1934) The question of the validity of the *N-T* tests has generally been ignored (Krey and Wesley, 1932, Kandel, 1936; Johnson, 1936, Robertson and Tryhorn, 1937), but in a few cases intelligent examination questions have been used as the criterion When a poor correlation was obtained between the *N-T* tests and the intelligent examination questions, which is the kind of correlation that is generally obtained, the apologists for the *N-T* tests have generally said that the so-called essay examination questions are unreliable, and that the *N-T* tests are better, although Tharp (1927), Cheydleur (1929), Sims (1933), Stalnaker and Stalnaker (1934a), Desclos (1936), and Stalnaker (1937) did not find that the so-called essay examinations were unreliable. One would expect that there would be some variation when different people grade intelligent examination papers, and there is also some variation when different people select different test items and include them in different intelligence tests

Some investigators have given a group of students true-false, multiple-choice, and completion tests, and so-called essay examination questions; and have then correlated each of the four scores with the average of the four scores on the four tests These correlations are often surprisingly high because each score is a common factor in the average of all four scores, but it makes very little difference what the correlations are when the so-called essay examination questions are used as the criterion

Various methods have been used to boost the low raw correlation reliabilities of the *N-T* tests, and after these somewhat questionable statistical operations, it has been claimed that the reliablity of the *N-T* tests is high because the "corrected" correlations are moderately high, and that the reliability of the so-called essay examinations is low because the raw correlations are moderately low. Some have thought that it was better statistical form to "correct" the reliability coefficients of the *N-T* tests than to "correct" the reliability coefficients of the so-called essay examinations, the statistical principles apparently operating differently in the two cases

Several experimenters have correlated the *N-T* tests and the so-called essay examinations with intelligence tests although the intelligence tests themselves may be made up of *N-T* test items

The chance factor in guessing at true-false and other *N-T* tests does not operate as it is often claimed, because tossing 40 or 50 coins once is by no means the same as tossing one coin 10,000 times. The factor of chance operates in some of the *N-T* tests in somewhat the same way that it operates in a hand of cards in the game of bridge It is difficult for any of the academic grades of an individual student to be valid or valuable when there is such a persistent and prominent factor of chance.

Kinney and Eurich's (1932) summary and Lee and Symonds' (1933, 1934) reviews of the *N-T* tests cover somewhat indefinite periods of time, and like many of the reviews of educational psychology, they seem to include principally "selected" references These reviews of the literature on the *N-T* tests are very vague on the question of the validity of the *N-T* tests, and the books which recommend experimental studies of the *N-T* tests also pass over the question of validity in an equally light manner Several of the "criteria" which seem to have been indirectly referred to in the rather hurried and forced discussions of the validity of the *N-T* tests have a strange appearance, and it seems that sufficient positive evidence has inadvertently accumulated to show that the *N-T* tests have a poor validity and generally do not measure what they are supposed to measure.

5. *Little Disagreements Among the Experts*

It has been claimed that there are several natural difficulties in the way of making out good *N-T* tests which are not always appreciated by those who experiment with them (Arnold, 1927, Worcester, 1928, Bird and Andrew, 1936, 1937, Paterson, 1936), and several of these difficulties depend upon the fact that an *N-T* question is not a good question if a large number of the students know the answer It is generally impossible to know whether or not an *N-T* test item will be diagnostic of anything before it is tried out, and after it has been tried out it cannot be used again The result of this situation is that those who work with the *N-T* tests are practically always experimenting, although some have used the chronic difficulty in making out the *N-T* tests as an argument in their favor

There are little disagreements among the experts on whether or not the students should be instructed to guess (Votaw, 1936), on the amount of time the students should be allowed in taking the *N-T* tests (Rogers, 1937), on how the items should be arranged in the *N-T* tests (Flanagan, 1937), whether or not the *N-T* tests should be scored by the students immediately after they take them (Cocks, 1929), whether or not the students should record their degree of confidence for each item of an *N-T* test (Soderquist, 1936), whether the score on the true-false test should be the number right or the number right minus the number wrong, whether or not true-false items should be weighted for importance in the scoring (Corey, 1930;

Potthoff and Barnett, 1932), and whether or not the students should give any reasons for their answers or guesses (McClusky and Curtis, 1926)

When the *N-T* test items are weighted for importance, even the so-called "objectivity" of scoring the *N-T* tests is destroyed

6 *The Influence of the New-Type Tests*

When the *N-T* tests are used extensively in a course, the students' attitudes tend to be unfavorable and they often learn very little They may not study for the *N-T* tests because, in spite of the instructor's arguments to the contrary, they feel that there is nothing to study for that is important and intelligent If they do study for the *N-T* tests and learn something, much of what they learn is not in an integrated form If they receive a low grade on an *N-T* test, it is impossible to show them the important facts and principles which they do not know What students say about the different types of examinations does not always have an entirely clear meaning, although this matter is often referred to in the literature, because the verbal attitudes of students who are untrained in psychological and statistical matters may be easily influenced by arguments in favor of a particular kind of examination, especially when the examination is very easy to take

Meyer (1934, 1935, 1936) has shown that college students learn more when they study for intelligent examination questions than when they study for true-false, multiple-choice, and completion tests, and it has also been found that those students who know the most about a subject are the students who object most strongly to the *N-T* tests (McClusky and Curtis, 1926, Fenton and Lehman, 1928, Tallmadge and Douglass, 1934).

7 *Cases of Specific Learning Where Formal Examinations are Not Necessary*

It may make little difference whether one uses a recall or a recognition test when the learning material is made up of nonsense syllables, and a fairly high correlation may be obtained between a multiple-choice and a completion vocabulary test with French words as memory material (Henmon, 1929, 1929a, Seibert, 1932) There are relatively few meaningful associations in the nonsense syllables, and the psychological processes of a high school or college student in learning an unfamiliar French word are somewhat similar to those of a grammar school child in learning to spell

In the case of a few specific forms of learning, a measure of efficiency may be obtained by the use of the *N-T* tests, but the difficulty in using this fact as an argument in favor of the *N-T* tests is that it is in just these cases of specific learning that formal tests and examinations are not necessary (Bardy, 1923) Some use may be made of the *N-T* tests with such subjects as arithmetic (Hawkes, 1936, 341-345) and spelling, but there are some arguments against giving any formal examinations in the elementary schools When groups of elementary school children are given impersonal *N-T* tests on, for example, arithmetic and spelling, perhaps in the hope of obtaining some kind of comparability, all of the disadvantages of the external examiner are present, and these disadvantages are serious in the elementary schools.

F. A BRIEF EXPERIMENT ON THE "OBJECTIVITY" OF THE NEW-TYPE TESTS

The principal purpose of the experiment described in the present section was to throw some light on the frequently made claim that the *N-T* tests are "objective."

We were using the intelligent-question method of teaching, as described in Section *C* above, with the 29 students in a class in experimental psychology, and the class was being given a written examination every three or four weeks The third examination was to include a selection from 26 already familiar questions on General Principles of Sensation, General Visceral Sensations, Kinesthetic Sensations, and Painful Sensations.

One week before this examination was to be given, the question was raised as to whether or not any members of the class would be willing to construct true-false tests on the same fairly definite material on which they were to be examined in the regular written examination, and several students volunteered to make up some true-false tests The members of the class understood that they would be graded on their answers to the intelligent or so-called essay examination questions. In this regular written examination, the members of the class were asked four questions about (*a*) the attributes of sensations, (*b*) sensory thresholds, (*c*) the nerves involved in general visceral sensations, and (*d*) the results of Jacobson's studies of progressive relaxation Four students out of the 29 did not take this third regular written examination because they were not prepared

At the next weekly laboratory meeting of the class, on the same or on the following day, the 25 students who had taken the regular written examination were given six different mimeographed true-false tests (on the same material covered by the written examination) which had been constructed by six members of the class The students who made out the true-false tests and the number of items in each test were as follows

Mary Egan	11 items	Marian Radke	25 items
Elsie Kohn	25 items	Jean Sutherland	25 items
Ruth Kowalke	21 items	Mary Thomson	13 items

The key used in scoring each true-false test was furnished by the student who constructed the test None of the subjects taking the true-false tests omitted any items, and we used the total number right as the score for each subject on each true-false test

The four students (Alvin Bruss, Lorena Cowgill, Edward Enders, and Rachel Stare) who did not take the regular written examination because they were unprepared were asked to read and grade the regular written examination papers, and another undergraduate student, Raymond Stone, who was somewhat out of touch with the material of the course, was also asked to read and grade these regular written examination papers None of the five readers had had any previous experience in grading papers, and none of them performed their 4-hour task with excessive enthusiasm. Each of the five readers read one question at a time through all of the papers, graded it on an absolute numerical scale from 0 to 13, and did not attempt to break up any of the topics into their "elements." The five readers were aware that some facts were much more important than others, and they attempted to give their grades with this in mind

Several correlations have been calculated using the grades and the scores of the 19 subjects who took both the regular written examination and the six true-false tests but who did not construct a true-false test themselves. The five students who served as readers did not take the regular written examination or the true-false tests.

Each of the five different readers gave each of the 19 subjects a total numerical grade on the regular written examination, and the correlations between the grades given by the five different readers

TABLE 1

CORRELATIONS BETWEEN THE GRADES GIVEN BY FIVE DIFFERENT READERS OF A SET OF INTELLIGENT OR SO-CALLED ESSAY EXAMINATION PAPERS

	Bruss	Cowgill	Enders	Stare	Stone	Mdn r
Bruss	—	+.62	+.64	+.55	+.81	+.63
Cowgill	+.62	—	+.52	+.36	+.35	+.44
Enders	+.64	+.52	—	+.79	+.49	+.58
Stare	+.55	+.36	+.79	—	+.49	+.52
Stone	+.81	+.35	+.49	+.49	—	+.49
Mdn r	+.63	+.44	+.58	+.52	+.49	

of the written examination are shown in Table 1. The 10 different correlations range from +.35 to +.81, with an Av r at +.56 and a Mdn r at +.53.

Each of six different students constructed a true-false test on the same material covered by the regular written examination, and the correlations between the scores on the six different true-false tests are shown in Table 2. The 15 different correlations range from —.15 to +.69, with an Av r at +.29 and a Mdn r at +.32.

TABLE 2

CORRELATIONS BETWEEN THE SCORES ON SIX DIFFERENT TRUE-FALSE TESTS WHICH COVERED THE SAME MATERIAL

	Egan	Kohn	Kowalke	Radke	Sutherland	Thomson	Mdn r
Egan	—	+.13	+.13	+.18	+.02	—.15	+.13
Kohn	+.13	—	+.33	+.25	+.49	+.38	+.33
Kowalke	+.13	+.33	—	+.62	+.37	+.10	+.33
Radke	+.18	+.25	+.62	—	+.53	+.32	+.32
Sutherland	+.02	+.49	+.37	+.53	—	+.69	+.49
Thomson	—.15	+.38	+.10	+.32	+.69	—	+.32
Mdn r	+.13	+.33	+.33	+.32	+.49	+.32	

These results show that the five different students who graded the answers to the intelligent or so-called essay examination questions agreed with each other to a greater extent than the six students who constructed the six different true-false tests; that is to say, the grading of the answers to the intelligent or so-called essay examination questions was more reliable than the construction of the true-false tests.

At this point I presume I should "correct" the low raw correlation reliability coefficients of the true-false tests, but at the same time accept the reliability coefficients of the readers of the written examination at their face value Or perhaps I should show that the new-type tests are all right by correlating new-type Test *A* with the average of new-type Tests *A*, *B*, and *C* In order to show that the new-type tests measure some fixed native "ability," perhaps I should correlate my new-type tests with an intelligence test which is itself partly made up of new-type test items. In the opinion of some, I should surely claim that, after all, good new-type tests are difficult to make out, and I should make the usual statement that really good new-type tests can be properly made out only by experts working a considerable period of time, I should also neglect to mention the fact that knowledge and experience improve the reliability of grading answers to intelligent examination questions, and finally, I should argue that the chronic difficulty in constructing good new-type tests only goes to show the practical value of these new-type tests. All of these things and arguments have frequently been done and made, but the reasons for doing these things and making these arguments are not clear

Each of the five different readers of the written examination gave each of the 19 subjects a numerical grade, and we have accepted the average of the five grades given each subject as the best and most sensible indication of what the subjects knew about the material, and these averages have therefore been used as the criterion The scores of the subjects on each of the six different true-false tests have been correlated with this criterion, and these correlations, showing that the validity of the true-false tests is poor, are as follows

Egan	+ 44
Kohn	+ 49
Kowalke	+ 50
Radke	+ 66
Sutherland	+ 47
Thomson	+ 16

At this point I should doubtless argue that these results, if properly "interpreted," mean that the new-type tests are good and valid tests, and that the reason the correlations with the criterion are not higher is that the so-called essay examinations are notoriously unreliable However, if anything can be found with which the new-type tests correlate reasonably well this should be looked upon as an excellent "criterion", but if nothing can be found with which the new-type tests correlate reasonably well, then it should be remarked that a suitable and proper criterion for the new-type tests is really very difficult to find The new-type tests are all right, the items in the new-type tests have been properly selected, and the trouble simply lies in the criterion which was not used in selecting the items. One should select the test items first, and attempt to find a suitable criterion afterwards

The procedure used in constructing many tests has been (*a*) to select the particular items for the test, (*b*) give the test a name and assume that the test is a good one, (*c*) say that no questions can be raised about the test because it is "objective," and (*d*) fail to mention that the particular items of the test had to be consciously selected and put in the test by the person who made up the test It has been found that if different people make out new-type tests on the same intelligent material, the new-type tests of the different people may correlate quite poorly with each other (Pullias, 1937). It is strange that the experiment of correlating different new-type tests *made out by different people* has not been more frequently carried out, especially since some of those who stress the practical value of the new-type tests are practically always experimenting

The construction of all examinations and the taking of all examinations are "subjective" processes inasmuch as the conscious-verbal processes of the teacher and the students are always involved. In the brief study that we have described, we have found that the construction of true-false tests was more "subjective" than the grading of students' answers to intelligent examination questions

G CONCLUDING STATEMENT

In the present discussion we have been principally concerned with the intelligent meaningful material of courses taught by the lecture-reading-discussion method, and such results as have been obtained by the use of material with almost as few associations as nonsense syllables are not applicable to our problem It seems in general that the traditional *N-T* tests should not be used in examining university, college, or high school students on any meaningful material, and that the more these tests are used in courses dealing with cultural material, the more these courses resemble the "over-pedagogicalized" textbooks "in whose bellies"—as William James once said—"little children's minds are turned to ashes."

The use of some kind of N-T tests which can be scored without error may seem to have an administrative advantage in the matter of comparability, but this advantage is somewhat illusory since the *N-T* tests generally do not measure what they are supposed to measure.

A well financed attempt was made to sell the American *N-T* tests to representatives of several foreign countries at the international conferences on examinations held at Eastbourne, England, in 1931 and at Folkestone, England, in 1935 (Monroe, 1931, 1936), and this conference is an interesting commentary on the problem of examinations. A few attempts were made to show that the *N-T* tests had some value, but most of these attempts were made at the beginning of the conference. The conference as a whole was opposed to the use of the *N-T* tests as a substitute for valid examinations, and the various reasons for rather flatly rejecting the *N-T* tests are stated in a most admirable fashion by Hartog (1931), Bouglé (1931, 1936), Sadler (1931), Burns (1931), Ballard (1936), and Desclos (1936a). Kandel's (1936) postmortem interpretation of the results of these international conferences did not improve the validity of the *N-T* tests nor decrease the general importance of the problem of examinations.

The *I-Q* method of teaching and testing which we have described is not offered as a general solution of all of our educational problems, but it is in harmony with several important psychological principles, and it could be used in connection with the lecture-reading-discussion part of many graduate and undergraduate courses. It appears that the general principles and spirit of our method would also have some practical advantage and saving grace in the elementary schools, junior high schools, and high schools, because in spite of all that has been said about democracy and education, it does seem that students should learn something before they come to college and should acquire some kind of cultural and intellectual information at all levels of our educational system.

References

1. Achilles, E M Experimental studies in recall and recognition *Arch of Psychol*, 1920, 6, No 44

2 Anastasi, A Further studies on the memory factor *Arch. of Psychol*, 1932, No 142

3 Arnold, H L. Analysis of discrepancies between true-false and simple recall examinations. *J Educ Psychol*, 1927, 18, 414-420.

4 Bagley, D A critical survey of objective estimates in the teaching of English *Brit J Educ Psychol*, 1937, 7, 57-71, 138-155

5 Ballard, P B In *Conference on Examinations* (Ed by P Monroe). *Columbia Univ, Teach Coll, Bur Pub*, 1936 (Pp 39-45)

6 Bardy, J An investigation of the written examination *Univ Penn, Ph D Thesis*, 1923

7 Bird, C, & Andrew, D M Concerning the length of new-type examinations *J Educ Psychol*, 1936, 27, 641-654

8 ——— The comparative validity of new-type questions *J Educ Psychol*, 1937, 28, 241-248

9 Bolton, E B The relation of memory to intelligence *J Exper Psychol.*, 1931, 14, 37-67

10 Bouglé, C In *Conference on Examinations* (Ed by P. Monroe) *Columbia Univ, Teach Coll, Bur Pub*, 1931 (Pp 38-50, 99-109)

11. ——— In *Conference on Examinations* (Ed by P. Monroe) *Columbia Univ, Teach Coll, Bur Pub*, 1936 (Pp 155-162)

12 Brinkley, S G Values of new type examinations in the high school *Teach Coll. Contrib Educ*, 1924, No 161

13 Brinkmeier, I. H, & Keys, N Circumstantiality as a factor in guessing on true-false examinations *J Educ Psychol*, 1930, 21, 681-694

14. Burns, C. D In *Conference on Examinations* (Ed by P Monroe) *Columbia Univ Teach Coll, Bur Pub*, 1931 (Pp. 226-234)

15 Cason, H The essay examination and the new type test *School & Soc*, 1931, 34, 413-418

16 ——— The learning and retention of pleasant and unpleasant activities *Arch of Psychol*, 1932, 21, No 134

17 Chen, H S *Psychol Abst*, 1933, 7, No 2105

18. Cheydleur, F D The relative reliability of the old and new type modern language examinations *French Rev*, 1929, 2

19 Claparède, E Récognition et moïité *Arch de Psychol*, 1911, 11, 79-90

20. Clarke, H M Recall and recognition for faces and names *J Appl Psychol*, 1934, 18, 757-763

21 Cocks, A W The Pedagogical Value of the True-False Examinations Baltimore Warwick & York, 1929

22 Corey, S M The effects of weighting exercises in a new type of examination *J Educ Psychol*, 1930, 21, 383-385

23 Desclos, A (Ed) La correction des épreuves écrites dans les examens Enquête experimentale sur le baccalauréate Paris Maison du Livre, 1936

24 ——— In *Conference on Examinations* (Ed by P Monroe) *Columbia Univ. Teach Coll Bur Pub*, 1936 (Pp 45-46, 131-155, 279-282, 292-293

25 English, H B, *et al* Studies in substance memorization *J Gen Psychol*, 1934, **11**, 233-260

26. Fenton, N, & Lehman, H C The true-false question and the student's sense of fairness *School & Soc*, 1928, **28**, 115-16

27 Flanagan, J C A proposed procedure for increasing the efficiency of objective tests *J Educ Psychol*, 1937, **28**, 17-21

28 Gamble, E A M, & Calkins, M W Über die Bedeutung von Wortvorstellungen fur die Unterscheidung von Qualitaten sukgessiver Reize *Zsch f Psychol*, 1903, **33**, 161-170

29 Gates, A I The Improvement of Reading New York Macmillan, 1927

30 Hall, C S Intercorrelations of measures of human learning *Psychol Rev*, 1936, **43**, 179-196

31 Hamilton, E R The Art of Interrogation New York Harcourt Brace, 1929

32 Hanawalt, N G Memory trace for figures in recall and recognition *Arch of Psychol*, 1937, No. 216

33 Hartog, Sir Philip In *Conference on Examinations* (Ed by P Monroe) *Columbia Univ Teach Coll, Bur Pub*, 1931 (Pp 30-36)

34 Hawkes, H E, *et al* The Construction and Use of Achievement Examinations New York Houghton Mifflin, 1936.

35 Heine, R Ueber Wiedererkennen und ruckwirkende Hemmung *Zsch f Psychol*, 1914, **68**, 161-236

36 Henmon, V A C Achievement Tests in the Modern Foreign Languages New York Macmillan 1929

37 ———, *et al* Prognostic Tests in the Modern Foreign Languages New York Macmillan, 1929

38 Hill, G E The effect of changed responses in true-false tests *J Educ Psychol*, 1937, **28**, 308-310

39 Hoffding, H Zur Theorie der Wiedererkennen *Phil Studien*, 1893, **8**, 86-96

40 Hollingworth, H L Characteristic differences between recall and recognition *Amer J Psychol*, 1913, **24**, 532-544

41 Husband, R W Intercorrelations between different forms of learning (In press)

42 Jenness, A The recognition of facial expressions of emotions *Psychol Bull*, 1932, **29**, 324-350.

43 Jersild, A T Examination as an aid to learning *J Educ Psychol*, 1929, **20**, 602-609

44 Johnson, P O The differential functions of examinations *J Educ Res*, 1936, **30**, 93-103

45 Jones, E S Comprehensive Examinations in American Colleges New York Ass Amer Colleges, 1933

46 Kandel, I L Examinations and their substitutes in the United States New York Carnegie Found Advance Teach, 1936

47 KATZAROFF, D Contribution à l'etude de la recognition. *Arch de Psychol*, 1911, **11**, 2-78

48 KELLEY, T L, & KREY, A C Tests and Measurements in the Social Sciences New York Scribners, 1934

49 KIMBALL, E P As for examinations *School & Soc*, 1928, **27**, 571-572

50. KINNEY, L B, & EURICH, A C A summary of investigations comparing different types of tests. *School & Soc*, 1932, **36**, 540-544

51 KREY, A C., & WESLEY, E B Does the new-type test measure results of instruction in the social studies? *Hist Outlook*, 1932, **23**, 7-21.

52 LEE, A L An experimental study of retention and its relation to intelligence *Psychol Monog*, 1925, **34**, No 157

53. LEE, J M, & SYMONDS, P M New-type or objective tests A summary of recent investigations *J Educ Psychol*, 1933, **24**, 21-38

54 ——— New-type or objective tests, A summary or recent investigations (October, 1931-October, 1933) *J. Educ Psychol*, 1934, **25**, 161-184

55. LEHMANN, A. Ueber Wiedererkennen *Phil Studien*, 1889, **5**, 96-156

56 LONG, J A., & SANDIFORD, P The validation of test items *Dept Educ Res., Univ Toronto Bull.*, 1935, No. 3.

57. MCCLUSKY, H Y, & CURTIS, F D A modified form of the true-false test *J. Educ Res.*, 1926, **14**, 213-224

58 MACCURDY, J T Common Principles in Psychology and Physiology Cambridge Cambridge Univ. Press, 1928

59 MEYER, G An experimental study of the old and new types of examination. I. The effect of the examination set on memory *J Educ Psychol*, 1934, **25**, 641-661

60. ——— An experimental study of the old and new types of examination· II. Methods of study *J Educ Psychol*, 1935, **26**, 30-40

61. ——— The effect on recall and recognition of the examination set in classroom situations *J Educ Psychol*, 1936, **27**, 81-99.

62 MONROE, P (Ed) *Conference on Examinations Columbia Univ Teach Coll, Bur Pub*, 1931, 1936

63 MYERS, G C A comparative study of recognition and recall *Psychol Rev*, 1914, **21**, 442-456

64 ODELL, C W Traditional Examinations and New-Type Tests New York Century, 1928

65 ORLEANS, J S, & SEALY, G A Objective Tests Yonkers-on-Hudson World Book, 1928.

66 OWEN, R B Recognition a logical and experimental study. *Psychol Monog.*, 1915, **20**, No 86

67 PATERSON, D G Preparation and Use of New-Type Examinations. Yonkers-on-Hudson World Book, 1927

68 ———, *et al* Revising test items to increase validity *J of Psychol*, 1936, **2**, 221-228

69. PIERON, H La critique expérimentale des methode d'examen. *Bull. Soc fr de péd*, 1928, **27**, 20-26

70 PILLSBURY, W. B The Fundamentals of Psychology New York. Macmillan, 1934

71 Potthoff, E F, & Barnett, N E A comparison of marks based upon weighted and unweighted items in a new-type examination. *J Educ Psychol*, 1932, **23**, 92-98

72 Pulias, E V Variability in Results from New-type Achievement Tests Durham, N C Duke Univ Press, 1937

73 Rabaud, E How Animals Find Their Way About A study of distant orientation and place-recognition New York Harcourt Brace, 1928

74 Robertson, R K, & Tryhorn, F G Objective test form in a school certificate examination *Brit J Educ Psychol*, 1937, **7**, 156-161

75 Rogers, H W The time factor in objective quizzes. *School & Soc*, 1937, **46**, 239-240.

76 Ruch, G M The objective or New-Type Examination New York Scott, Foresman, 1929

77 Sadler, Sir Michael In *Conference on Examinations* (Ed by P Monroe) *Columbia Univ, Teach Coll, Bur Pub*, 1931 (Pp 122-130)

78 Seibert, L C A series of experiments on the learning of French vocabulary *Johns Hopkins Univ Stud Educ*, 1932, No 18

79 Seward, G H Recognition time as a measure of confidence *Arch of Psychol*, 1928, **16**, No 99

80 Sims, V M Objective tests and teachers' measurements *School & Soc*, 1932, **36**, 300-302

81 ——— Improving the measuring qualities of an essay examination *J Educ Res*, 1933, **27**, 20-31.

82 Skaggs, E B Some critical comments on certain prevailing concepts and methods used in mental testing *J Appl Psychol*, 1927, **11**, 503-508

83 Soderquist, H O A new method of weighting scores in a true-false test *J Educ. Res*, 1936, **30**, 290-292

84 Stalnaker, J M Essay examinations reliably read *School & Soc*, 1937, **46**, 671-672

85 Stalnaker, J M, & Stalnaker, R C Open-book examinations *J Higher Educ*, 1934, **5**, 117-120

86 ——— Reliable reading of essay tests *School Rev*, 1934, **42**, 599-605

87 Strong, M H, & Strong, E K The nature of recognition memory and of the localization of recognitions *Amer J Psychol*, 1916, **27**, 341-362

88 Symonds, P M Size of recognition and recall vocabularies *School & Soc*, 1926, **24**, 559-560

89 Tallmadge, M, & Douglass, H R What college students believe they think about certain types of examinations *School & Soc*, 1934, **39**, 349-352

90. Terman, L M, & Merrill, M A Measuring Intelligence A guide to the administration of the new revised Stanford-Binet tests of intelligence New York Houghton Mifflin, 1937

91 Tharp, J B The new examination versus the old in foreign languages *School & Soc*, 1927, **26**, 691-694.

92 Votaw, D F The effect of do-not-guess directions upon the validity of true-false or multiple choice tests *J Educ Psychol*, 1936, **27**, 698-703

93. Welborn, E L, & English, H Logical learning and retention. *Psychol Bull*, 1937, **34**, 1-20

94. Wheat, H G The Teaching of Reading Boston Ginn, 1923

95 Whipple, G. M Manual of Mental and Physical Tests II Complex Processes Baltimore Warwick & York, 1915

96 Wolfe, H K Untersuchung uber das Tongedachtniss *Phil Stud*, 1886, **3**, 534-571

97 Woods, E L An experimental analysis of the process of recognizing *Amer J Psychol*, 1915, **26**, 313-387

98 Woodworth, R S Psychology New York Holt, 1929

99. Worcester, D A Prevalent errors in new-type examinations *J Educ Res*, 1928, **18**, 48-52

100 Wylie, M An experimental study of recognition and recall in abnormal mental cases *Psychol Monog*, 1930, **39**, No 180

Department of Psychology
University of Wisconsin
Madison, Wisconsin

The Journal of Genetic Psychology, 1939, 54, 391-402

THE PERCEIVING AND THINKING OF NORMAL AND SUBNORMAL ADOLESCENTS AND CHILDREN ON A SIMPLE DRAWING TASK*

New York City Trade Extension Classes

Theodora M. Abel and Jane B. Sill

A. Introductory

Rupp has shown some of the difficulties encountered by mental defectives in carrying out drawing patterns that present no manual obstacle (3). Due to deficiency in both differentiating and integrating wholes and parts, the feeble-minded individual may, for instance, copy a diamond as a simple oval, or draw a circle and add small pointed corners along its circumference to represent angles. Also children perform differently from adults in carrying out drawing tasks. Like mental defectives, they may show either too little differentiation, as in drawing one circle to represent both the body and head of a man, or inadequate differentiation, as in giving a man two eyes in a profile drawing. Piaget has called these two modes of behavior, "syncretism" and "juxtaposition" respectively, both due to inability to form sufficient syntheses in the perceptual and thinking functions (2).

The purpose of this study was to compare performances on the same task of subnormal and normal subjects of equivalent chronological age, and of normal and superior children, younger chronologically, but with mental ages fairly comparable to those of the subnormals, in order to determine specifically the similarities and differences in performance among these different developmental groups. The division of a square into smaller ones was mentioned by Lewin in his discussion of a dynamic theory of the feeble-minded (1, p. 220). He reported that the feeble-minded solved the task by making "rows of smaller squares" within the large square, rather than by dividing the square by drawing "parallel lines through the entire figure." He did not indicate, however, at what level of mental deficiency this type of performance took place, nor what were the

---

*Received in the Editorial Office on June 27, 1938.

other possible modes of performance either for the feeble-minded or the normal individual, he was merely illustrating an intellectual defect of the feeble-minded This task seemed to be simple and concise, and one where the results could be analyzed and recorded by fairly objective procedures.

In presenting this task to a few subnormal subjects in a preliminary series, we found we had inadvertently used two different methods of instruction, either we said to divide the square *into squares* or to divide it into *smaller squares*. It appeared that these preliminary results varied according to the type of instruction used In fact, one subject asked what we meant by "smaller"—did we mean "smaller and smaller"? We decided then to give the two methods of instruction in order to include in our problem the question of the influence of type of verbal instruction upon performance, controlling the conditions to the extent of changing only one word in the instructions.

B PROCEDURE

1 *Task*

The *S*s were presented individually with a drawing of a square outline (4" x 4") One half of the *S*s were told to take a pencil and divide the square "into squares" (Method *A*). The other half were told to divide the square "into smaller squares" (Method *B*) All *S*s were allowed to work as long as they wanted and to do whatever seemed best to them The experiment was carried out during 1937 and 1938

2 *Subjects*

The subjects were 322 adolescents and 108 children distributed into the following groups [1]

> I Subnormal adolescent school group 208 girls with *IQ* 50-79 on the Otis Self-Administering Intermediate Examination, *CA* 14-17, 86% 15 and 16

---

[1] The *S*s in Groups I and II were in the Manhattan Industrial High School for Girls and its Extension Classes, in New York City; the children in Group III were in the Barnard School for Girls, the Chapin School, and Friends Seminary, New York City, the *S*s in Group IV were in Letchworth Village, a New York State institution for the mentally deficient Thanks are due the principals of the above schools and the research staff of Letchworth Village for their cooperation in making this study possible

II Normal adolescent school group 94 girls with *IQ* 80-110 on verbal group tests, *CA* 14-17, 79% 15 and 16

III. Normal and superior children 91 girls and 17 boys, *IQ* on Stanford-Binet 92-146, 26% between 92 and 110 74% over 100 *CA* 5-9

IV Institutionalized subnormal adolescent group 11 boys and 9 girls, *IQ* on Stanford-Binet 40-49, *C.A* 16-19

C RESULTS

The various ways in which the task was carried out fell quite distinctly into six categories as follows

1 *"X-section"* the square as a whole divided into squares by unbroken parallel lines running through the entire square, horizontally and vertically.
2 *"Nest"* concentric squares producing a camera or nest effect
3 *"Bricks"* built up by successive additions of squares to the first small one drawn adjoining sides of original square, result resembling #1
4 *"Boxes"* squares drawn separated from, and within, the lines forming the original square, not in nest fashion.
5 *"Wrong"* square divided, but not into anything resembling a square, or boxes or circles drawn all over the paper
6 *"Mixed"* two or more methods used, one of them usually wrong (Category 5)

1 *Groups I and II*

Our first comparison is between the performances of Groups I and II, the subnormal (*IQ* 50-79) and normal (*IQ* 80-110) adolescents In Table 1 are shown the results for these two groups

TABLE 1

Groups	*IQ*	Method	Total No	X-section		Nest		Bricks		Boxes		Mixed		Wrong	
				N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
1. Subnormal	50-79	A	104	54	52	26	25	9	09	1	01	2	02	12	11
		B	104	29	28	50	48	11	10	5	05	4	04	5	05
II Normal	80-110	A	47	34	73	5	11	3	06	2	04	0		3	06
		B	47	34	73	6	.12	5	11	1	02	1	02	0	

including the frequency of responses in each category, under the two methods of instruction, *A* and *B*.

In the subnormal group, the cross-section and the nest categories contain the large majority of responses. Under Method *A*, one-half of the responses fell under cross-section and one-fourth under nest. Under Method *B*, however, this situation is reversed, one-fourth falling under cross-section and one-half under nest. The association of performance with the method of instruction used, *A* or *B*, is found to be significant statistically by the method of X^2. (Table 2).

TABLE 2
SUBNORMAL GROUP

Method	*N* X-section	*N* Nest	Total *N*.		
A	54	26	80	X^2	15.10
B	29	50	79	P	< .01
Total *N*.	83	76	159		

In the normal group, on the other hand, the cross-section category alone contains the large majority of responses. Under both Methods *A* and *B* three-fourths of the performances are cross-sections. It would appear then, that for the subnormal group, the addition of the word "smaller" to the instructions induces a subject, who might otherwise cross-section the square, to divide it into progressively smaller squares, each one a smaller replica of the one before, while in the normal group, the word "smaller" is *not* used with the same connotation, "progressively smaller." Regardless of the verbal instruction, the majority of the normal subjects interpret the task as one of dividing the whole square into squares by cross-section.

The difference in mode of performance between Method *A* and

TABLE 3

IQ	Method	Total No.	X-Section		Nest	
			N	%	*N*	%
70-79	A	30	17	57	6	20
	B	30	7	24	18	60
80-89	A	24	15	63	4	17
	B	24	16	67	5	21

(These *IQ* ranges correspond roughly to mental age ranges of 10-6—11-10 and 12-0—13-4, respectively.)

B appears between the *IQ* ranges of 70-79 and 80-89 as can be seen from an examination of Table 3

In Method *B*, the subjects with 70-79 *IQ* make many more nests than they do by Method *A* For the subjects with 80-89 *IQ* there is no change in performance under Method *B* Furthermore, the results for the *IQ* range 90-110 are the same as in the given 80-89 range On the other hand, the difference between the two methods of instruction in the 70-79 *IQ* range is equally large in the ranges 60-69 and 50-59 Apparently, between the *IQ* ranges 70-79 and 80-89 (as measured by the Otis test) there is a shift in the thinking processes of the *S*s with a resultant change in the meaning attached to the directions given.

As we already had scores on a non-language test (Pintner Non-Language Mental Test) for 241 cases in these adolescent groups, we were interested in finding out whether there was a definite level of non-verbal intelligence where a shift in performance by the *B* method took place as it did on the Otis, a strictly verbal test Among these subjects this non-language *IQ* range was from 53 through 130 We divided the subjects into two *IQ* groups as we had on the Otis Their performance on the square by the classifications of cross-section and nest is indicated in Table 4 (The other four categories are excluded here as the frequencies were quite small and similar by the two methods)

TABLE 4

Pintner Non-Language *IQ*	Method	Total No	X-section *N*	X-section %	Nest *N*	Nest %			
50-79	A	53	21	40	14	26	X^2	5 28	05 > P > 02
	B	52	13	25	26	50			
80-130	A	67	43	64	15	22	X^2	6 64	P < 01
	B	69	29	43	28	41			

Here there is a significant relation between type of response and type of instruction in both the lower and higher *IQ* groups. At both levels there is an increase in nests and decrease in cross-section under Method *B* Even with the higher Pintner *IQ*s (90 up to 130) there is still a marked difference in performance by Methods *A* and *B* A difference in verbal intelligence (as measured by the

Otis *IQ*) rather than non-verbal intelligence is the dominant factor in producing the difference in comprehension resulting in dividing the square into squares or into "smaller" squares.

These results suggest possibilities for further research on the understanding of language among those subnormals whose verbal ability falls below a definite level. Do they take all comparative adjectives and adverbs as signifying progression rather than as a difference in degree? For instance, if they were presented with a sample block painted red, with several other unpainted blocks and with some painting material, and were told to paint the blocks redder than the sample given, would they make each block progressively redder than the last one? These results also point to the importance of keeping verbal instructions constant in testing procedures if comparable results are to be expected.

To return to our initial results, we want to point out also that a few individuals in both the subnormal and normal groups made responses falling into our other four categories, bricks, boxes, mixed and wrong, 49, or 24 per cent of the subnormals and 15, or 19 per cent of the normals, performed in this manner. It is interesting to note the overlapping in performance of the two different developmental groups, subnormal and normal, and that, irrespective of the intelligence level, some *S*s performed the task in a less syncretic and less integrated manner (making bricks or boxes) while others solved the problem with a high degree of synthesis of the whole. Even among the lower *IQ* levels of subnormality (*IQ* 50-59) 32 per cent gave cross-section responses, almost as many as in the higher levels (*IQ* 70-79), where 49 per cent gave cross-section responses.

2 *Group III*

The results for the children of normal and superior intelligence, by the two methods of instruction *A* and *B* are given in Table 5.

TABLE 5

Children	Method	Total	X-section		Nest		Bricks		Boxes		Wrong		Mixed	
		N	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%
CA 5-9														
IQ 92-146	A	54	24	44	15	28	6	11	6	11	2	04	1	02
	B	54	26	48	13	24	3	06	5	09	4	07	3	06

The children do not react differently under Methods *A* and *B*, the frequency of their responses in the different categories being practically identical under the different types of instruction Consequently, we may say that as a group they do not respond specifically to the word "smaller" as did the adolescent subnormals In this respect their behavior is more like that of the normal adolescents. On the other hand, the children cross-section the square only two-thirds as frequently as do the normal adolescents (under both Methods *A* and *B*, only about 50 per cent of the children and 75 per cent of the normal adolescents did so). Their frequency of response by cross-section is like that of the subnormal adolescents under Method *A*, but is not reduced under Method *B* as it was for the subnormals. This correspondence of response type under Method *B* with classification of our subjects into children and subnormal adolescents is significant by the method of X^2 as shown in Table 6.

TABLE 6

By Method *B*	X-section *N*	Nest *N*	Total *N*		
Children	26	13	39	X^2	9 146
Subnormal	29	50	79	P	$<$ 01
Total *N*	55	63	118		

At the same time, the children make twice as many nest responses as the normal adolescents (26 per cent as against 12 per cent by Methods *A* and *B* combined) and equally as many as the subnormals by Method *A* (25 per cent), but this frequency of nest responses, as we have said, is not increased under Method *B* as it was for the subnormals (24 per cent as against 48 per cent)

Only 27 per cent of the children's responses by both methods of instruction, *A* and *B*, fall in the other categories, bricks, boxes, mixed and wrong Their behavior in this respect is like that of the subnormals (24 per cent) and only slightly different from that of the normal adolescents (19 per cent).

We next compared the performance of the children divided first into younger and older chronological age groups (*CA* 5-0—6-11 and *CA* 7-0—9-11), and secondly, into the younger and older mental age groups (*MA* 6-0—8-11 and *MA* 7-0—9-11), to see whether there were differences in modes of response under the two methods of instruction as chronological age and mental age were varied. We found no differences. Consequently, we grouped together the results of Methods *A* and *B*, and compared the responses of the younger and older chronological and mental age groups. The results for the chronological age levels are given in Table 7

TABLE 7

Chronological ages	Total *N*	X-section *N*	%	Nest *N*	%	Bricks *N*	%	Boxes *N*	%	Mixed *N*	%	Wrong *N*	%
5-0—6-11	46	14	30	15	32	4	09	5	11	4	09	4	09
7-0—9-11	62	36	58	13	21	6	10	6	10	1	01	0	

As chronological age increases the number of whole responses of the cross-section category increases appreciably, but, concomitantly, the frequency of nest responses does not decrease appreciably as it did in comparing the subnormal and normal adolescents. Among the children, the differences in response between the five and six-years olds and the seven through nine-year olds lies in a change of frequency of responses in the other categories, bricks, boxes, etc , as well as the cross-section category

Among the younger children, *CA* 5-0—6-11, 37 per cent make responses falling in the four categories, bricks, boxes, mixed and wrong, while among the older children *CA* 7-0—9-11, only 21 per cent respond in this manner

Our next step was to compare the frequency of responses falling in the cross-section category with those falling in the other five categories grouped together in order to find out the relationship between the chronological age level and mode of performance The results are (Methods *A* and *B* combined) given in Table 8

TABLE 8

Chronological age	X-section *N*	Other five categories *N*	Total *N*		
5-0—6-11	14	32	46	X^2	8 108
7-0—9-11	36	26	62	P	< 01
Total *N*	50	58	108		

Using the X^2 method we see that there is a significant connection between mode of response (cross-section and all other categories) and chronological age level

We carried out the same procedure for two different mental age levels among the children, grouping together the responses under Methods *A* and *B* with the results given in Table 9.

TABLE 9

Mental age	Total *N*	X-section *N*	 %	Nest *N*	 %	Bricks *N*	 %	Boxes *N*	 %	Mixed *N*	 %	Wrong *N*	 %
6-0—8-11	62	22	36	21	34	5	08	6	10	4	06	4	06
9-0—12-11	46	27	59	8	17	5	11	5	11	1	02	0	

As can be seen, there is a definite increase in frequency of cross-section responses with increase in mental age as was the case with the increase in chronological age. But also, there is a decided decrease in the frequency of nest responses as mental age rises, which was not the case with increase in chronological age. By the X^2 method, the relationship between mental age level and mode of response (cross-section and nest) we obtain a X^2 of 5.576, $.02 > P > .01$, showing that these results are statistically reliable.

As in the case of the chronological age comparisons, there is also a decrease in the frequency of responses in the four other categories (bricks, boxes, mixed and wrong) as mental age increases. In the 6-8 mental age group, 31 per cent of the total responses fell in these categories, while in the 9-12 mental age group, only 24 per cent of the responses do so. The latter frequency is the same as the frequency of these modes of response found among the subnormal adolescents whose mental ages were roughly equivalent to those of these older children. But as we have seen, there is one qualitative difference in verbal comprehension between the children and the subnormal adolescents. In comprehending the total sentence of the instructions, the children did not distinguish the word "smaller" as having a specific meaning of progression, as did the subnormals. We do not mean to imply that the children necessarily responded in a more mature manner than the subnormals in comprehending the two sets of instructions as having the same meaning of division into smaller parts. The higher grade subnormals may well have been making finer distinctions of meaning than did the children.

3. *Group IV*

The results for our lower subnormal group, *IQ* 40-49, cannot be compared directly with the other groups as to methods of instruction *A* and *B*, since the subjects had difficulty understanding the task required and frequently needed to be given instructions by both methods in order to obtain a response It appeared that if they failed to perform under Method *A*, the addition of Method *B*, or the word "smaller," induced them to respond Sometimes they merely made an additional square outside the square outline and had to be given the further instruction of dividing the square inside the square The distribution of their responses is given in Table 10

TABLE 10

Group IV	Method	Total *N*	X-section *N* %	Nest *N* %	Bricks *N* %	Boxes *N* %	Wrong *N* %	Mixed *N* %
IQ 40-49	A and B	20	1 05	3 15	0	11 .55	4 20	1 05

Here, the majority of responses fall in the boxes category, only one *S* among the 20 was able to divide the square as a whole in a cross-section Consequently, we can see that there is a considerable difference in the comprehension of the task between this lower level of mental deficiency and our other group of subnormals of greater intelligence The group with the *IQ* range 40-49 performed more frequently in the manner suggested by Lewin (1, p 220).

Although the mental ages of the imbeciles overlap those of the children in the mental age range 6-8, they do not succeed as frequently as did these children, in performing in an integrated manner, either cross-sectioning the square as a whole or even dividing up the whole into nests of squares Thirty-six per cent of the children with mental ages 6-8 cross-sectioned the square as a whole, and 34 per cent made nests of squares, while only 10 per cent made discrete boxes.

We must realize, of course, that the imbeciles were institutionalized, with little chance for educational opportunities, while the children had cultural opportunities for development both in their homes and in private school. Over and above this difference, however, is the difference in quality of thinking, the children showing greater flexibility, the adult imbeciles greater rigidity or inflexibility in mental functions Lewin has explained this difference between the child and the subnormal as being due to a difference in ability to think in a concrete or in an abstract manner "This difficulty concerning abstract thinking," he says, "appears pronounced in the feeble-minded at an age at which the child of equivalent mental age no longer shows this peculiar concreteness" (1, p 223). Certainly, the cross-sectioning of the square as a whole or making a graded series of squares one within the other (nests) requires greater abstraction in thinking than does making discrete squares

4 *Individual Differences*

Within each of our categories of responses, we find further individual differences in performances Sometimes in using the bricks method the final result appeared practically identical with that of the cross-section, while at other times the sides of the bricks were separated from the square outline or one column was separated from the others One might say that in the former response, an inefficient cross-section, or more integrated method was used, while in the latter, an efficient boxes, or less integrated method was used Responses might become either more or less integrated as the performance continued A camera response sometimes became a square spiral or even started as such. Perhaps this indicated a more complete surrender to the idea of "smaller" rather than "divide" or "square"

The length of time spent on the task depended not only on the type of the response itself, but also on the general mode of attack, whether hesitantly, impulsively or carefully There were individuals who divided the square as exactly as possible, into equal quarters, while others started on one side and drew parallel lines, with no apparent measurement ahead of time Sometimes a strong perseverative trend was displayed in that the square might be cross-sectioned for as many as 25 minutes, until a gray shading covered the whole square,—or the concentric squares of a camera were drawn very close together and continued to a dot

As in many test performances, the questions and remarks showed certain individual differences in number and kind Some subjects asked many questions before starting, while others began immediately Questions were asked about the directions, the purpose of the task and the other individuals' performances Puzzlement at both the simplicity of the task and the lack of a time limit were expressed by some, as well as worry, amusement, or unconcern with the whole activity. At times, insight into the thought processes of an individual were gained from remarks and questions.

Final responses, therefore, might be identical as to category, yet vary in the method of developing the response, in perseveration, or in the general approach to the problem.

D. Conclusion

In this study we found quantitative and qualitative differences in modes of performance on a simple drawing task among different developmental groups, children, adolescent subnormals (imbeciles and higher grades), and normal adolescents. The most striking differences found were, in the first place, the greater concreteness of performance on the part of the lower grade subnormals in contrast to the young children of equivalent mental age, and secondly, the special mode of comprehension of the verbal instructions on the part of the higher grade adolescent subnormals. This mode of comprehension was not only qualitatively different from the normal adolescents of equivalent chronological ages, but also qualitatively different from the younger children of equivalent mental age.

On the other hand, we found a certain amount of overlapping in modes of performance among the different developmental groups, even on a simple task, suggesting that a rigid classification of groups on the basis of chronological age or of composite score on a standardized test (mental age) does not tell us everything about the mental functioning of the individual, the ways in which he perceives, comprehends, organizes his ideas, and carries out the tasks he has to do. The normal individual does not invariably react in a highly integrated and synthetic manner to even a simple task requirement, nor does the child and subnormal react necessarily in a less organized and more syncretic manner to the same task requirement. And finally, the child and the older subnormal of equivalent mental age do not necessarily react in the same manner to the same task situation.

References

1. Lewin, K. A Dynamic Theory of Personality. (Translated by D. K. Adams and K. E. Zener.) New York: McGraw-Hill, 1935.

2. Piaget, P. J. Le Jugement et le Raisonnement chez L'Enfant. Paris: Alcan, 1924.

3. Rupp, H. Über optische Analyse. *Psychol. Forsch.*, 1923, 4, 262-300.

11 West 11 Street
New York City

The Journal of Genetic Psychology, 1939, 54, 403-424.

NATIONALITY PREFERENCES AND STEREOTYPES OF COLORED CHILDREN*

Psychological Service Center, St. Louis

H. Meltzer

A. The Problem

Made in America are the feeling and thinking reactions of children born and raised in the United States. So made are the nationality and race preferences and stereotypes of colored children. A competent anthropologist in a study of the social origin of the negro goes so far as to assert that the psychology of the negro in the United States is such that "the sanctions of the white population are the accepted sanctions of negroes," that "white behavior patterns are not only automatically adhered to but consciously striven for" (3, p. 253). To what extent is this tendency to accept the cultural values of the majority group manifested in the nationality preferences and attitudes of colored children? Few, if any, of the previous studies of negroes throw much light on this problem. This topic, for example, is not at all touched upon in the comprehensive volume on *The American Negro* (14) published by the American Academy of Political and Social Science. Illustrations of prejudices and general discussions of the social significance of race prejudice are found in books on social psychology[1] and in such books as Donald Young's *American Minority People* (15) or Lasker's *Race Attitudes in Children* (5). Bogardus (1) reported a social distance study on a group of 200 negroes in his volume on *Immigration and Race Attitudes*. But no systematic study of this particular problem, the race preferences and stereotypes of colored children, has been made. The present study represents an attempt to fill this gap.[2] The purpose is thus to invesitgate the nature of the race and nationality attitudes of colored children by analyzing the preferences expressed by a representative group, comparing these preferences with those expressed by a group of white children, and analyzing the reasons given by them for the preferences they make Specifically the questions investigated in this study are: What nationalities or races do colored children prefer? How do these preferences differ from those made by a group of white children? How intense are the feeling reactions towards nations and races they like or dislike? What reasons do they give for their expressed feelings? To what extent does stereotypy characterize these reasons? What is the nature of their race and nationality stereotypes? What do these stereotypes signify as "sets" for social interaction in the American scene?

---

*Received in the Editorial Office on July 5, 1938.

[1] See, for example, Chapter XI in (9) and Chapters XVIII and XIX in (16).

[2] The writer is indebted to many individuals for help which made possible the bringing of the present study to completion. His first indebtedness is to Mrs. Myra F. Dubinsky, who was particularly helpful in preparing the study for publication. Others who have helped with the scoring or statistical work are Mr. Ernest Haveman, Mrs. Alice E. Hellman, and a number of *F.E.R.A.* workers sent by Dean F. M. Debatin of Washington University. For cooperation of school authorities the writer is indebted to

B. THE CHILDREN EXAMINED

Three hundred and sixty-four colored children selected from one of the better schools of the colored district in St Louis were the subjects of this experiment St. Louis, it will be remembered, is dominantly Southern in the emotional atmosphere it sets for the negro in the city. In the district where this school is located 14 per cent of the population over 10 years of age were illiterate in 1930 Only 100 of the group of 4000 illiterate negroes, however, were under 21. In 1932 only 8 per cent of the dwellings in this district were occupied by the owners. Of the houses in the area, 32 per cent had a value over $3000 or a rental over $30 a month The median rental of the houses was about $26 50 per month. The children were selected from grades five to eight inclusive The age range was from nine to sixteen The average age of the children from the fifth grade was 11 5, of the sixth grade children, 12 3, of the seventh grade children, 13 5, and of the children from the eighth grade, 13 8 This is somewhat, but not very much, above the age-grade norms for the population as a whole. The average educational retardation of the group would be somewhat more than one year The group of white children used for comparing preferences of the colored group was composed of 1265 children ranging in grades from five to eight inclusive, and selected from three eco-

---

the following Dr Henry Gerling, Superintendent of Schools, Mr George R Johnson, Director of Tests and Measurements; and Miss Mathilde C. Gecks, Assistant Superintendent To all teachers and children who cooperated the writer expresses his gratitude nomic level neighborhoods.[3] The examining was done in 1934.

Some light on the cultural background of the colored group can be obtained from a survey of the books and magazines these children read. Ninety-seven of the 364 children (26.6 per cent) reported reading no books whatever. Only 18 reported using library books. School books received frequent mention. Thirteen reported reading fairy tales, 21, history books, 13, the Bible. Seven said they read *Treasure Island.* Three reported reading stories of different nations. *Tom Sawyer* received two mentions, adventure books received seven mentions, poetry, eight, the *Book of Knowledge,* four, and *Sport Stories,* two. Among the seventy-some books which received one mention are *The Prince and the Pauper*, *Salt and Pepper*, *Alice in Wonderland*, *Life of Booker T. Washington*, *Negroes in the World War*, *Things for Girls to Do*, *Young Trailers*, and *Huckleberry Finn*.

One hundred and ninety-two of the 364 children (52.7 per cent) reported that they read no magazines. *True Story* outranked all other magazines and received 39 mentions. Movie magazines received 20 mentions, *Saturday Evening Post,* 19, *Liberty,* 14, *Love Mirror,* 13, *Literary Digest,* 12, *Detective Stories,* 10. Between five and ten mentions were given to the following: *American, True Romance, True Confessions, Time, Ladies Home Journal, Woman's Home Companion, Child Life, National Geographic, McCalls, Popular Science,* and *Boy's Life*. Among the magazines that received from one to five mentions are *Red Book, Collier's, Delineator, Mystery, Boy Scout Magazine, Western Magazine, Pictorial Review,* and *American Youth.*

C. PROCEDURE

The original plan was to study the problems of nationality preferences by the use of three supplementary methods: (*a*) the Paired Comparison Method as described by Thurstone,[4] for its statistical advantages, (*b*) a form devised by the writer to get information concerning the degree as well as direction of feeling for each nationality or race and statement of reasons for feeling expressed, (*c*) individual interviewing procedure ranging from free association to direct questioning The present study is limited to a consideration of the findings from the use of the first two methods.

---

[3] For description of the three economic level areas see Meltzer (8, p. 592), and also the description of subject and procedure in (7).

[4] For discussion of this method the reader is referred to the following articles: Thurstone (13), Guilford (2) and Meltzer (7).

The instructions for the paired comparison sheet read as follows

> This is an experimental study of group attitudes You are asked merely to underline the one of each pair with whom you would rather associate For example, the first pair is
>
> Englishman - South American
>
> If, in general, you prefer to associate with Englishmen rather than with South Americans, underline Englishman If you prefer, in general, to associate with South Americans, underline South American If you find it difficult to decide for any pair simply underline one of them anyway Be sure to underline one of each pair even if you have to guess

The nations and races included are American, Armenian, Chinaman, Englishman, Frenchman, German, Greek, Hindu, Irishman, Italian, Japanese, Jew, Mexican, Negro, Pole, Russian, Scotchman, Spaniard, South American, Swede, and Turk.

On the second mimeographed sheet the children were instructed to check the degree of feeling towards each of the 21 nations and races listed. The range of intensity of feeling included were intense like, like, neutral, dislike, and intense dislike After checking their feeling reactions they were then instructed to go back and on a space to the right, write out the reason for each of the feeling reactions expressed toward each nation or race Care was taken that the instructions were clearly understood and ample time was given the children for completing these forms

D NATIONALITY PREFERENCES OF COLORED AND WHITE CHILDREN

The order of preference for nations and races expressed by colored children as well as the preferences expressed by a comparable group of white children is given in Table 1

The foregoing are the results of the paired comparison method of 364 colored children and 1265 white children previously described In the first two columns to the right of the list of nations are reported the Thurstone scale values for the two groups of children How the different nations and race are ranked by the colored children as compared with the white children is here indicated The last column

TABLE 1

NATIONALITY PREFERENCES OF COLORED AND WHITE CHILDREN

Nation or race		Scale values 364 Colored children	 1265 White children	Rank for White	Difference in rank order
1	Negro	0.0000	3.6841	20	+19
2	American	0.4438	0.0000	1	— 1
3	English	1.0306	0.6840	2	— 1
4.	French	1.6181	1.2051	3	— 1
5	South American	1.6223	1.9794	8	+ 3
6	Mexican	1.6935	2.1381	11	+ 5
7	Spaniard	1.9104	1.8922	6	— 1
8	Italian	2.1736	1.9983	9	— 1
9	Irish	2.2743	1.2668	4	— 5
10	Japanese	2.2938	2.8263	17	+ 7
11	Scotch	2.3005	1.9438	7	— 4
12	German	2.3710	1.8540	5	— 7
13	Russian	2.6114	2.4088	12	— 1
14	Chinaman	2.6141	3.1490	18	+ 4
15	Swede	2.6960	2.0769	10	— 5
16	Armenian	2.7679	2.8983	16	0
17.	Jew	2.8574	2.5438	13	— 4
18	Greek	3.0353	2.8977	15	— 3
19	Pole	3.3730	2.7720	14	— 5
20	Turk	3.3777	3.4833	19	— 1
21	Hindu	3.7832	3.7410	21	0

in the table shows the difference in the rank order of the groups. The largest difference in rank is in the placement of the negro. Ranked twentieth by the white children, he is ranked first by the negro children. Other nations and races that are markedly more favored by the colored than by the white are the Japanese, Mexican, South American and Chinaman. Less favored by the colored children to the extent indicated in the last column in Table 1 are the following nations or races: German, Pole, Swede, Jew, Scotch, and Greek.

The correlation between the rank order of the colored children and the white children is .60 ± .09. This relatively low correlation is largely due to the fact that the colored children change their own rank from twentieth to first, and also markedly change the rank order of the nations and races which colored children think of as having greater possibilities for friendly interacting relations than do white children. Roughly this correlation of .60 is a measure of the extent to which the nationality preference patterns of colored children are completely American. The variability of the colored group is actually somewhat larger than that of the white According to Thurstone

> "a wide range of scale values indicates rather strong and rather uniform national preferences An internationally minded group would give proportions hovering more closely about 50 and hence the scale separations would be, on the whole, smaller than the scale separations for a prejudiced group"

Guilford argues that the chief factor determining variability is the heterogeneity of the group. In the light of the facts presented here it seems doubtful that a smaller variability is in this instance a reliable measure of greater tolerance. A more plausible explanation is that the colored group of children have, in addition to the general American cultural pattern, an internal solidarity that is responsible for their reacting differently to those nations and races that they feel more closely related to And it is this more markedly favorable reaction to groups which they feel closer to and more marked unfavorable reactions to groups which have shown more signs of rejecting them which explains their somewhat larger variability In their nationality preferences, therefore, colored children do not show evidence of the undiscriminating imitations of white patterns that Herskovits gives as a frequently found negro characteristic Instead their preference patterns might be more accurately said to be the end-product of the interplay of more commonly expressed sentiments of Americans in general, and their preferences made in the light of their own internal solidarity as a minority group The correlation of .60 may be taken as a rough measure of the extent to which their preference patterns are in accord with the generally accepted public race sentiments in the United States

E. INTENSITY OF FEELING REACTIONS

The intensity of feeling reactions to each of the 21 nations and races expressed by the 364 colored children are given in Table 2

A few facts about differences in the degree of feeling are apparent from the data in the foregoing table First, they express many more intense like reactions than intense dislike reactions—27.4 per cent like as compared to 8 6 per cent dislike, many more like reactions than dislike—24.3 per cent as against 10 2 per cent Nations or races for which they express an intense dislike are Jew, Turk,

TABLE 2
INTENSITY OF FEELING REACTIONS TO 21 NATIONS AND RACES EXPRESSED BY COLORED CHILDREN

Nation or race	Intense like	Like	Neutral	Dislike	Intense dislike	Omissions	Valence score
Negro	88.7	7.7	1.9	0.8	0.0	0.8	1.86
American	84.8	12.4	0.3	0.8	0.8	0.8	1.81
Englishman	43.9	33.4	14.1	4.1	2.5	1.9	1.18
South American	31.5	36.2	24.3	3.6	3.9	0.6	.88
Mexican	32.2	31.8	24.0	6.4	3.0	2.5	.86
Frenchman	30.4	32.9	23.8	7.7	3.3	1.9	.81
Spaniard	30.4	27.9	27.6	8.3	3.6	2.2	.75
Japanese	27.6	30.9	22.9	10.2	5.8	2.5	.66
Italian	23.8	29.0	27.1	11.3	5.2	3.6	.57
Irishman	21.8	27.1	31.8	9.7	6.4	3.3	.50
Scotchman	18.5	31.8	32.3	9.4	5.8	2.2	.46
Swede	17.4	27.3	35.9	9.9	7.2	2.2	.39
Chinaman	21.0	30.1	22.4	13.5	11.1	1.9	.37
Russian	15.7	27.6	37.8	10.2	6.6	1.9	.36
Armenian	13.0	18.8	42.8	12.2	8.3	5.0	.17
Greek	13.5	22.9	32.6	16.6	12.7	1.7	.08
Pole	9.4	19.9	44.5	11.6	9.9	4.7	.08
German	15.7	23.3	26.5	18.0	14.9	1.7	.07
Turk	8.0	16.3	39.0	16.9	16.6	3.3	—.13
Jew	17.7	15.2	24.0	14.6	25.7	2.8	—.16
Hindu	9.4	9.1	33.7	18.5	26.8	2.5	—.45
Total	27.4	24.3	27.1	10.2	8.6	2.4	

German, Greek, Chinaman, and the Pole. Much less often reported are intense dislikes for other nations and races. The 10 nations or races that they express a marked intense liking for are, in the order of preference: Negro, American, Englishman, South American, Mexican, Frenchman, Spaniard, Japanese, Italian, and Irishman. The most neutral reactions in the order listed are given to the following nations or races: Pole, Armenian, Turk, Russian, and Swede. The valence score reported in the last column of Table 2 is obtained by assigning a numerical value to each degree of feeling expressed, obtaining the algebraic sum of all values, and dividing the number by the total number of children reacting. The scores assigned for each degree of feeling were: +2 for intense like and —2 for intense dislike, +1 for like and —1 for disliske, and 0 for neutral responses. The specific percentages of all degrees of feeling for each nation can be read directly from Table 2. The correlation between the ranks based on the valence score so obtained and the ranks based on the scale values obtained by the paired comparison method is .95±.01. This tends to affirm the conclusion of Saffir (12) to the effect that choice among psycho-physical methods may be made on the basis of a practical convenience, rather than relative validity.

F. Race and Nationality Concepts of Colored Children

So much for the preference and intensity of the feeling reactions colored children have for nations and races herein investigated. What of their reasons for these feelings? What notions of the various nations and races do they have? To what extent do these tend to be stereotyped?

In racial stereotypes we have, according to Katz and Braly (4, p. 280) conditioned responses of varying degrees of aversion or acceptance toward racial labels "which when they can be applied to individuals as they can in the Negro because of skin color, we respond to him not as a human being but as a personification of the symbol we have learned to despise." The results of our comparative study of preferences suggest that their preferences are of American origin modified by what might be referred to as a Negro patterning in the American scene. What of the nature of their stereotypes? Braly and Katz in their study of the racial stereotypes of 100 college students had their subjects check a list of traits for describing each of 10 groups studied. The 10 groups were: German, Italian, Negro, Irish, English, Jew, American, Chinese, Japanese, and Turk. In their study they report the 12 traits checked for each of the 10 groups. Stereotypy they indicate by reporting the least number of traits which must be taken to include 50 per cent of the possible assignment for each group.

Zeligs (17) in her study of race concepts of 200 sixth grade children reports only the single most common concept given by the children to each of the 38 nations and races studied. Illustrative examples of concepts given for favorable, neutral, and unfavorable reactions to Americans, Jews, and Negroes are also included by her.

No quantitative treatment of the data is made by this investigator. The method used in this study is purposed to yield the advantage of naturalness and spontaneity obtained by Zeligs who had the children give their own notions rather than check a list given to them by the investigator, and also the advantage of more quantitative treatment of the results made by Katz and Braly.

The reasons given by all of the 364 children for the feeling reactions they expressed toward each nation or race were classified and ranked. Space makes the inclusion of all the notions of all groups given by the children prohibitory. Instead will be reported the six most frequent notions of each nation or race the colored children expressed. These notions, the number of times each has been mentioned and the percentage for each frequency are all reported in Table 3.

TABLE 3

The Six Most Common Race and Nationality Notions of Colored Children

Nation or race	Six most common notions for each nation	Number of mentions	Per-cent
American	I am an American, it's my country	234	64.29
	They are O.K., nice, kind, good, fine	46	12.63
	I just like them	14	3.84
	They're my friends, acquaintances	12	3.30
	No reason given	9	2.47
	They give us jobs	6	1.65
	Total	321	88.18
Armenian	Don't know them, no contacts	241	58.80
	Nothing against them	30	8.25
	They're all right, nice, good, kind	26	7.13
	Just don't like them	18	4.95
	They are Americans	13	3.57
	I just like them	10	2.75
	Total	311	85.44
Chinaman	They are O.K., nice, kind, good, fine	53	14.55
	They're evil, mean, cruel, bad	30	8.25
	They're good cooks, I like their chop suey	24	6.59
	Don't know them, no contacts	22	6.05
	I just like them	19	5.23
	I have Chinese friends	15	4.12
	Total	163	44.78
Englishman	I can understand their language	67	18.42
	No reason given, no contacts	56	15.38
	They're O.K., nice, kind, good, fine	52	14.30
	They're like us, it's our mother country	42	11.50
	I just like them	16	4.40
	I have English friends, acquaintances	16	4.40
	Total	249	68.40
Frenchman	I don't know them, no reason	96	26.37
	They're O.K., nice, kind, good, fine	63	17.30
	I like their accent, language	32	8.80
	They're our allies	20	5.50
	Nothing against them	13	3.57
	Don't like them	12	3.30
	Total	236	64.83

TABLE 3 (*continued*)

Nation or race	Six most common notions for each nation	Number of mentions	Per-cent
German	They're warlike	58	15.91
	They're O.K., good, nice, kind, friendly	39	10.71
	I don't know them, no contacts	37	10.16
	I just like them	32	8.79
	They help when in need	27	7.42
	They talk funny	12	3.30
	Total	205	56.32
Greek	Don't know them, no contact	65	17.85
	They're O.K., nice, kind, good, fine	45	12.33
	They're good cooks	43	11.80
	Nothing against them	25	6.87
	They're evil, mean, cruel, bad	22	6.05
	Just don't like them	20	5.50
	Total	220	60.44
Hindu	Don't know them, no reason	149	40.93
	I don't like them	29	7.96
	They're evil, mean, bad, cruel	26	7.14
	Nothing against them	19	5.22
	They're O.K., nice, kind, good, fine	15	4.12
	They're sneaky, tricky	13	3.57
	Total	251	68.95
Irish	Don't know them, no reason	108	29.67
	They're nice, good, kind, O.K., fine	60	16.47
	Nothing against them	22	6.05
	Just don't like them	16	4.40
	They're evil, mean, bad, cruel	14	3.84
	I have Irish friends	14	3.84
	Total	234	64.28
Italian	Don't know them, no reason	94	25.82
	They're nice, kind, good, fine	58	15.91
	I have Italian friends	22	6.05
	Nothing against them	16	4.40
	Just don't like them	15	4.12
	They're evil, mean, bad, cruel	14	3.84
	Total	219	60.16
Japanese	Don't know them, no contact	70	19.23
	They're O.K., nice, good, kind, fine	54	14.83
	They're smart, clever, wise	22	6.05
	Appearance—dress	21	5.77
	Nothing against them	16	4.40
	They're warlike	13	3.57
	Total	196	53.85
Jew	They cheat, are unfair, dishonest	58	15.91
	They're nice, good, kind, fine, friendly	38	10.43
	I have Jewish friends	21	5.77
	No reason given	19	5.23
	They're evil, mean, bad, cruel	19	5.23
	Can't understand and dislike language	15	4.12
	Total	170	46.70

TABLE 3 (*continued*)

Nation or race	Six most common notions for each nation	Number of mentions	Per-cent
Mexican	Don't know them no reason	68	18.68
	They're O.K., nice, good, kind, fine	53	14.55
	I have Mexican friends	18	4.95
	Nothing against them	18	4.95
	I like their music and songs	16	4.40
	Just like them	13	3.57
	Total	186	51.09
Negro	I am a Negro, it's my race	270	74.17
	They're O.K., nice, good, kind, fine	22	6.05
	Just like them	11	3.02
	I have Negro friends	10	2.74
	No reason given	10	2.74
	It's the dark race, our color	8	2.17
	Total	331	90.90
Pole	Don't know them, no reason	197	54.12
	Nothing against them	24	6.59
	They're O.K., nice, good, kind, fine	22	6.05
	It's cold there, too far north	15	4.12
	Just don't like them	14	3.84
	Just like them	11	3.02
	Total	283	77.74
Russian	Don't know them, no reason	132	36.24
	They're O.K., nice, good, kind, friendly	48	13.19
	Nothing against them	20	5.50
	Appearance—dress	15	4.12
	They're industrious, ambitious	14	3.84
	Just don't like them	14	3.84
	Total	243	66.73
Scotchman	Don't know them, no reason	141	38.73
	They're O.K., nice, good, kind, fine	53	14.56
	Just like them	15	4.12
	Appearance—dress	13	3.57
	Customs	10	2.75
	Just don't like them	10	2.75
	Total	242	66.48
Spaniard	Don't know them, no reason	110	30.21
	They're O.K., nice, good, kind, friendly	35	9.61
	I like their dances	33	9.06
	I like their music and songs	20	5.50
	Just like them	15	4.12
	Appearance—dress	11	3.05
	Total	224	61.53
South American	Don't know them, no reason	110	30.21
	They're O.K., nice, good, kind, friendly	46	12.63
	It's near America	27	7.40
	Just like us	22	6.05
	I studied about them	22	6.05
	They raise good crops	18	4.97
	Total	245	67.31

TABLE 3 (*continued*)

Nation or race	Six most common notions for each nation	Number of mentions	Per-cent
Swede	Don't know them, no reason	155	42.58
	They're O.K., nice, good, kind, friendly	39	10.71
	Nothing against them	24	6.59
	Just like them	17	4.67
	Don't like them	17	4.67
	They're industrious, hard working	11	3.05
	Total	263	72.27
Turk	Don't know them, no reason	154	42.30
	They're warlike	35	9.61
	They're evil, bad, unfriendly	27	7.40
	Nothing against them	22	6.05
	They're O.K., nice, good, kind, friendly	22	6.05
	Just don't like them	18	4.97
	Total	278	76.38

A study of colored children's notions of Americans in part reported in Table 3 clearly suggests that their reasons for feeling as they do about them is due to their identification with the group. *I am an American* or *it is my country* are illustrative notions. The dependence of the negro on the whites for jobs is a notion that six of the young subjects already have acquired. Influenced by their study of American history are their reasons for feelings they express towards the Englishman. Illustrative are such reasons as, they are like us, it is our mother country, they are discoverers and explorers, they started the Industrial Revolution. Also influenced by history teaching are the following reasons for feeling reactions to Spaniards: are explorers and discoverers, they sank the Maine, are fond of gold. They were our allies in the Great War is a reason given for liking the Englishman, Frenchman, and Italian.

They're O.K., fine, nice, good, appears as one of the six most frequent reasons for the feelings expressed towards all of the 21 nations and races. They're evil, bad, mean, cruel, is given as one of the six most frequent notions for seven of the twenty-one nations and races. Nations or races so described include Chinaman, Greek, Hindu, Irish, Italian, Jew, and Turk. A liking for their music and songs is one of the six most frequent reasons given by the colored children for their feelings towards two groups—Spaniards and Mexicans. Industriousness as one of the six most common reasons for liking is expressed by the colored children towards only two groups—Russian and Swede.

Notions of the Chinaman that appear to be more in accord with the American stereotypes than are the six most frequently given include the following are peaceful, clean, honest, trustworthy, make beautiful dishes Other reactions to the Chinaman are I'm afraid of them, have long finger nails, are awkward, kill and chop up people, are superstitious, and they put a wall around them The colored children's notions of the Greeks seem to be influenced by fragmentary knowledge of early Greek history as well as modern Greece Representative reactions are: they are beautiful, are good athletes, their soldiers wear dresses, I like to see them run. Characteristic notions of the Scotch include such reactions as I like their bagpipes, I like their accent, they hate the Irish, their men wear dresses, are quarrelsome.

They are warlike is the reason given for 118 dislikes, 35 neutral reactions, and 16 likes Of all the nations studied Germany is by far most often so described (58 mentions). Turkey is second with 35 mentions Japan is a poor third with 13 mentions Other nations so described are China, 9 mentions; Russia, 8, Spain, 7, France, 6 Italy and Mexico are each so described five times, England and Greece obtain four such mentions, and Ireland three Armenia, India, Poland, South America, and Sweden each obtain two such mentions The Negro and the Scotchman are each so described once

Peaceful appears as a reason for feeling reactions in the following nations or races two mentions for Chinaman, Frenchman, and Swede The Englishman, German, Greek, Irish, Japanese, Mexican, Pole, and Russian each receive one such mention.

G THE NATURE AND EXTENT OF STEREOTYPY IN COLORED CHILDREN'S RACE AND NATIONALITY CONCEPTS

That children's race concepts and values tend to be stereotypes—pictures in the head—in the sense which Lippman[5] first analyzed their influence in the formation of public opinion can be easily observed from the data reported in Table 3 A more concise picture of the extent of stereotypy in reasons given can be obtained from the data reported in Table 4 Here is reported the convergence of judgment as indicated by the concentration of answers found in the reactions of the children to each nation or race. Concentration or convergence is reported in terms of percentage of the total number of possibilities for the first six most common notions. The columns are presented in reverse order, the total per cent for the first six notions appears in the first column and the per cent of the most frequently mentioned appears in the last column

---

[5] For a discussion of the meaning of stereotypes the reader is referred to Walter Lippman's *Public Opinion*" (6), Stuart Rice's "*Quantitative Methods in Politics*" (10) and "*Personality and the Social Group*" (11), and Meltzer's "*Personification of Ideals and Stereotypes in Problem Children*" (8)

The presence of a large degree of stereotypy in the convergence or concentration of answers reported in Table 4 is easily observable.

TABLE 4
STEREOTYPY OF COLORED CHILDREN'S RACE CONCEPTS AS INDICATED BY CONVERGENCE OF JUDGMENT IN THEIR SIX MOST FREQUENTLY GIVEN NOTIONS

Nation or race	Total of 1st 6	Total of 1st 5	Total of 1st 4	Total of 1st 3	Total of 1st 2	Total of 1st only
American	88.18	86.53	84.06	80.76	76.92	64.29
Armenian	85.44	82.69	79.12	74.17	67.04	58.80
Chinaman	44.78	40.66	35.43	29.38	22.79	14.55
Englishman	68.40	64.00	59.60	48.10	33.80	18.42
Frenchman	64.83	61.53	57.96	52.46	43.67	26.37
German	56.32	52.99	45.57	36.78	26.62	15.91
Greek	60.44	54.90	48.85	41.98	30.18	17.85
Hindu	68.95	65.37	61.25	56.03	48.89	40.93
Irishman	64.28	60.43	56.59	52.19	46.14	29.67
Italian	60.16	56.32	52.18	47.78	41.73	25.82
Japanese	53.85	50.28	45.88	40.11	34.06	19.23
Jew	46.70	42.57	37.34	32.11	26.34	15.91
Mexican	51.09	47.53	43.13	38.18	33.23	18.68
Negro	90.90	88.72	85.98	83.24	80.22	74.17
Pole	77.74	74.72	70.88	66.76	60.71	54.12
Russian	66.73	62.90	59.05	54.93	49.43	36.24
Scotchman	66.48	63.73	60.98	57.41	53.29	38.73
Spaniard	61.53	58.50	54.38	48.88	39.82	30.21
South American	67.31	62.34	56.29	50.24	42.84	30.21
Swede	72.27	69.22	64.55	59.88	53.29	42.58
Turk	76.38	71.41	65.36	59.31	51.91	42.30
Median	66.48	62.34	57.96	52.46	43.67	30.21
Range	(44.78–90.90)	(40.66–88.72)	(35.43–85.98)	(29.38–83.24)	(22.79–80.22)	(14.55–74.17)

The six most common notions include on the average 66.48 per cent of all the reasons given. In only two of the 21 possibilities do the six most common notions add up to less than 50 per cent of the total reasons given The median per cent of all reasons given included by the five most common notions is 62.34, for the first four is 57.96, for the first three is 52.46; the first two, 43.67, and the first only, 30.21. The range for the single most common reaction is from 14.55 to 74.17, for the first two most common reactions is from 22.79 to 80.22, for the first three is 29.38 to 83.24; the first four, 35.43 to 85.98, the first five, 40.66 to 88.72, and the first six, 44.78 for the Chinaman to 90.90 for the negro. That ignorance as well as preconceived pictures in the head contribute towards a convergence of responses is indicated by the large extent of concentration in reasons given for feelings expressed towards the Armenian, the least well known of nations and races studied. Exceeded only by the Negro and American is the degree of concentration of reasons given for feelings expressed towards the Armenian.

Further light on the nature as well as the extent of stereotypy

TABLE 5

STEREOTYPY AND CONSISTENCY IN REASONS FOR FEELING REACTIONS OF COLORED CHILDREN

Nation or race	Number of reasons for each feeling response Ratio like reason	 Ratio neutral reason	 Ratio dislike reason	% Reasons inconsistent with feelings
American	13.3	0	1.0	0.3
Armenian	6.9	14.2	9.5	5.2
Chinaman	4.3	3.2	3.3	11.9
Englishman	7.7	3.1	2.8	4.4
Frenchman	5.3	5.4	4.0	6.9
German	3.9	5.4	5.4	11.0
Greek	4.6	4.7	5.0	8.8
Hindu	3.0	6.7	6.4	5.5
Irishman	5.3	5.7	3.9	11.3
Italian	4.8	5.0	3.1	7.2
Japanese	4.9	4.0	3.6	6.9
Jew	4.2	4.3	6.2	11.3
Mexican	5.0	3.4	2.5	8.3
Negro	15.9	1.6	2.0	0.6
Pole	3.8	7.8	7.2	11.0
Russian	4.2	7.1	3.8	8.6
Scotchman	5.4	7.3	4.3	6.1
Spaniard	5.3	4.8	4.1	6.4
South American	6.4	4.7	2.8	2.2
Swede	4.2	8.0	5.3	6.6
Turk	3.6	10.1	8.1	9.4
Total	28.0	20.6	21.6	7.1

in reasons given by colored children for their feeling reactions to 21 nations or races can be obtained by a study of the data reported in Table 5 In the first column are presented the number of reasons for each favorable (like) reaction expressed In the second and third columns similar information is given for neutral and dislike reactions Reported in the last column in terms of per cent is the amount of inconsistency found in reasons given for feelings expressed

The most stereotypy in the like reactions of colored children is found in their reactions to the following nations or races the Hindu, the Turk, the Pole, the German, the Russian, the Jew, and the Swede The least stereotyped are first, the Negro, second, the American, and then a poor third is the Englishman Most stereotypy in the neutral feelings expressed by the colored children is in their reactions to the following nations or races American, Negro, Englishman, Chinaman, and Mexican The most stereotypy in the dislike reactions are the American, Negro, Mexican, Englishman, and South American The most inconsistency in feelings expressed and reasons given is found in the Chinaman The Irishman and the Jew tie for second Tied for third are the German and the Pole. In the order named, most consistency is found in responses to the American, the Negro, the South American, and the Englishman—all relatively very favored groups.

Further insight on the nature of the stereotypy found can be obtained by studying the differences in the concentration of reasons given for favorable, unfavorable, or neutral feeling reactions Such information is indicated in Table 6

TABLE 6
DIFFERENCES IN MEAN NUMBER OF REASONS GIVEN FOR LIKES, INDIFFERENCES AND DISLIKES EXPRESSED

	Mean	Standard Deviation	Standard Error of the Mean
Like	33 95	8 54	1 86
Indifference	17 57	9 78	2 02
Dislike	14 29	8 22	1 69

The foregoing data indicates that the number of reasons given for liking nations or races is larger in extent than are the number of reasons for expressing indifferent reactions or reactions of dislike. That is to say, the least stereotypy is found in the reasons given for favorable reactions and the most stereotypy is found in the reasons given for unfavorable reactions.

H. NATIONALITY AND RACE ATTITUDES OF COLORED CHILDREN

The social significance of an attitude lies in the fact that it represents a "set" or determining tendency for interaction. The extent to which colored children are set for various types of interacting adjustments with the 21 nations or races is indicated in Table 7.

TABLE 7

NATIONALITY AND RACE ATTITUDES OF COLORED CHILDREN

Nation or race	Identification % (in group)	% Friendly interacting reactions	% Neutral acceptance reactions	% Hostile interacting reactions	% Don't know and no reason given
American	64.6	31.8	0.0	0.8	2.8
Armenian	0.3	19.4	13.1	6.0	61.1
Chinaman	0.3	49.4	18.2	22.3	9.8
Englishman	3.4	68.6	8.6	3.4	15.6
Frenchman	0.0	56.9	11.7	4.7	26.7
German	0.0	35.1	17.3	28.4	19.2
Greek	0.0	31.0	18.4	20.4	30.2
Hindu	0.6	14.0	14.2	29.6	41.6
Irishman	0.8	41.8	18.4	8.8	30.2
Italian	0.0	46.3	15.4	11.8	26.7
Japanese	0.0	54.3	14.6	11.8	19.3
Jew	0.0	31.2	20.2	39.2	9.4
Mexican	1.2	56.5	15.6	7.2	19.6
Negro	74.3	20.7	0.8	0.8	3.6
Pole	0.3	19.1	15.9	8.1	56.8
Russian	0.0	33.9	17.2	10.9	37.9
Scotchman	0.0	38.9	15.2	6.2	39.7
Spaniard	0.3	50.3	12.1	5.7	31.3
South American	2.5	54.6	9.1	3.3	30.5
Swede	0.0	36.5	14.2	5.1	44.1
Turk	0.0	19.1	16.9	21.6	42.5
Average	7.2	38.6	13.6	12.2	28.4

Categories used for classifying styles of adjustment are first, *identification* (in group), second, *friendly interacting reactions*, third, *reactions which indicate neutral acceptance*, and fourth, *hostile interacting reactions*. The per cent of "don't know" answers reported in the last column are included here because they also indicate an unpredictable "set" for interaction. As a minority group with their own distinctive sensitivities no reactions are given by the colored children which would indicate a completely forbidden touch relationship with any group. Identification responses are made almost exclusively to American and Negro. Aside from these two groups the Englishman gets 13 such mentions, the South American, 9, the Mexican, 4. Friendly interacting reactions give a range from 19.1 per cent for the Pole and the Turk to 68 per cent for the Englishman. Nations or races which rate large percentages of neutral acceptance reactions are: Jew, 20.2 per cent or 71 mentions, Greek, 18.4 per cent or 66 mentions, Irishman, 18.4 per cent or 66 mentions, Chinaman, 18.2 per cent or 65 mentions, German, 17.3 per cent or 62 mentions, and Russian, 17.2 per cent or 60 mentions. Nations or races which rate large percentages of hostile interacting responses are: Jew, 39.2 per cent; Hindu, 29.6 per cent, German, 28.4 per cent, Chinaman, 22.3 per cent, and Turk, 21.6 per cent.

Of the population studied the colored children can get along in friendly manner or better with 45.8 per cent, can get along with a neutral acceptance or better with 59.6 per cent, towards only 12.2 per cent of the population do they have hostile interacting reactions. Twenty-eight and four-tenths per cent "don't know" answers are unpredictable in the lump, for that is given as reason for liking and disliking as well as neutral reactions.

I. SUMMARY

1. *Summary of Procedure*

Three hundred and sixty-four colored children, ranging in age from nine to sixteen and in grade from fourth to eighth, inclusive, were examined for nationality preferences, concepts and attitudes by the use of two supplementary methods: (*a*) Paired Comparison Method, (*b*) a form devised by the writer to obtain information concerning the degree as well as the direction of feeling for each of twenty-one nations or races and a statement of reasons for the feelings expressed. Used for comparative study were 1265 white children from three schools representing three economic level neighborhoods.[6]

---

[6] For more detailed description of this group see Meltzer (8, p. 592). Also an article on "*Group Differences in Nationality and Race Preferences of Children*" (7).

2. *Summary of Findings*

a. Nationality preferences of colored and white children. The order of preference for the 21 nations and races studied of the 364 colored children beginning with the most preferred were as follows: Negro, American, Englishman, Frenchman, South American, Mexican, Spaniard, Italian, Irishman, Japanese, Scotchman, German, Russian, Chinaman, Swede, Armenian, Jew, Greek, Pole, Turk, and Hindu. The scale values and rank order of the colored children as well as the white are indicated in Table 1. The largest difference in rank is in the placement of the negro. Ranked twentieth by the white children, he is ranked first by the negro children. Other nations and races that are markedly more favored by the colored than by the whites are Japanese, Mexican, South American, and Chinaman. Less favored by the colored children to the extent indicated in the last column in Table 1 are the following nations or races: German, Pole, Swede, Jew, Scotch, and Greek. The correlation between the rank orders of the white and colored children is .60. This correlation is interpreted as a rough measure of the extent to which the nationality preference patterns are completely American. Largely because of their distinctive differences in nations favored, their variability is larger than that of the white children. Variability is not, following Thurstone's suggestion, accepted as a measure of tolerance or international mindedness in a minority group like the negroes. Nor is it, as suggested by Guilford, an index of heterogeneity. It is rather the result of the interplay between their internal solidarity in the light of which they make preferences that characterize them as a distinct minority group and American accepted preference patterns in the light of which they express the more common or public preferences of Americans in general. Their greater preference for South Americans and Mexicans illustrate the former, their high ranking of the Englishman an illustration of the latter.

b. Intensity of feeling reactions. Toward the 21 nations or races studied, colored children express many more intense like reactions than dislike, many more like than neutral reactions. The percentage of the various degrees of feeling reactions expressed toward each of the 21 nations or races is reported in Table 2.

c. Race and nationality concepts of colored children. Influenced by their study of American history are many of the children's notions of the English and the Spaniards Among the reasons given for feeling reactions to the English, for example, are They are like us, it is our mother country, they are discoverers and explorers, they started the Industrial Revolution Among such reasons for feeling reactions to the Spaniards are They sank the Maine; they are fond of gold, and they are explorers and discoverers They were our allies in the Great War is a reason given for liking the English, French, and Italian

They are warlike is the reason given for 118 dislikes, 35 neutral reactions, and 16 likes Of all the nations studied, Germany is by far the most often so described (58 mentions) Turkey is second with 35 mentions Japan is a poor third with 13 mentions Other nations so described are China, 9 mentions, Russia, 8 mentions, Spain, 7, France, 6 Italy and Mexico are each so described five times, England and Greece obtain four such mentions, and Ireland, three Armenia, India, Poland, South America, and Sweden each obtain two such mentions The Negro and the Scotchman are each so described once. The six most frequently mentioned concepts for all of the 21 nations or races are reported in Table 3

d The nature and extent of stereotypy in colored children's race and nationality concepts Markedly stereotyped are the colored children's nationality and race concepts The extent is in the paper reported in two ways First, by the extent of convergence or concentration of answers The percentages are indicated in Table 3 In Table 4 stereotypy is reported by indicating the number of reasons for each feeling reaction, and further insight on the nature of the stereotypes is indicated by the per cent of reasons which are inconsistent with feelings. Illustrative facts indicating the extent of stereotypy are: The six most common notions include on the average 66.48 per cent of all the reasons given The median per cent of all reasons given included by the five most common notions is 62.34, for the first four is 57.96, for the first three is 52 46, for the first two, 43.67; and the first only, 30 21 Illustrative facts about the nature of the stereotypes are: Ignorance as well as preconceived pictures in the head contribute towards convergence of responsiveness or stereotypy Most consistency between feelings and reasons for race and nationality preferences is found in the children's reactions to the American, Negro, South American, and Englishman—all relatively favored groups The number of reasons for liking nations or races is larger in extent than the number of reasons for indifferent reactions or reactions of dislike Least stereotypy, that is, the least stereotyping is found in the reasons given for favorable reactions, and most stereotypy is found in the reasons given for unfavorable reactions

e Nationality and race attitudes of colored children All the reasons given when classified for styles of possible interaction fall into the categories indicated in Table 7—Identification, Friendly Reactions, Neutral Acceptance, Hostile Reactions, Don't Know answers Of the population studied the colored children can get along in friendly manner or better with 45.8 per cent, can get along with an attitude of neutral acceptance or better with 59 6 per cent Towards only 12 2 per cent of the population do they have hostile interacting reactions Twenty-eight and four-tenths per cent "don't know" answers are unpredictable in the lump, for that is given as reason for liking and disliking as well as neutral reactions The extent of the different "sets" for interaction towards each nation or race is reported in Table 7.

REFERENCES

1 BOGARDUS, E S Immigration and Race Attitudes New York, Heath, 1928 Pp 286

2 GUILFORD, J P Racial preferences of a thousand American university students *J Soc Psychol*, 1931, 2, 179-204

3 HERSKOVITS, M J Social history of the Negro Ch 7 in *A Handbook of Social Psychology*, edited by Carl Murchison Worcester, Mass Clark Univ Press, 1935 Pp 1195 (Pp 207-267)

4 KATZ, D & BRALY, K Racial stereotypes of one hundred college students *J Abn & Soc Psychol*, 1933, 28, 280-290

5 LASKER, B Race Attitudes in Children New York Holt, 1927 Pp 394

6 LIPPMAN, W Public Opinion New York Harcourt Brace, 1922 Pp 427

7 MELTZER, H Group differences in nationality and race preferences of children *Sociometry*, 1939, 2, 86-105

8 ——— Personification of ideals and stereotypes in problem children *Amer J Orthopsychiat*, 1932, 2, 384 399

9 MURPHY, G, & MURPHY, LOIS B Experimental Social Psychology, New York Harper, 1931 Pp 709

10 RICE, S A Quantitative Methods in Politics New York Crofts, 1928 Pp 331

11. ——— Stereotypes Ch 15 in *Personality and the Social Group,* edited by Ernest W Burgess Chicago Univ Chicago Press, 1929 Pp 230 (Pp 192-197.)

12 SAFFIR, M A comparative study of scales constructed by three psychophysical methods (Reported at the American Psychological Association meetings, 1937)

13. THURSTONE, L L An experimental study of nationality preferences. *J Gen Psychol*, 1928, **1**, 405-423

14 YOUNG, D (Editor) The American Negro Amer Acad Pol & Soc Sci, 1928, **140**, No 229 Pp 359

15 ——— American Minority Peoples. New York: Harper, 1932 Pp. 621

16 YOUNG, K Source Book for Social Psychology New York, Crofts, 1931 Pp 844 (Chapters 18 & 19)

17 ZELIGS, R, & HENDRICKSON, G Racial attitudes of two hundred sixth grade children *Soc & Soc Research*, 1933, **18**, 26-36

Psychological Service Center
4510 Maryland Avenue
St. Louis, Missouri

The Journal of Genetic Psychology, 1939, 54, 425-434

QUANTIFIED PRESSURE STIMULATION AND THE SPECIFICITY AND GENERALITY OF RESPONSE IN FETAL LIFE*

Psychology Departments of Tufts College and University of Rochester

LEONARD CARMICHAEL AND M. F. SMITH

In a study recently reported by Carmichael and Lehner (4), it was shown that in the guinea pig fetus temperature stimulation released patterns of behavior during almost the entire span of motile fetal life. It was further shown that the greater the difference between the temperature of the stimulus and the physiological zero of the organism, that is, the more intense the stimulus, the greater the relative number of responses released by the stimulus. The results of this investigation of temperature sensitivity in fetal life seemed to have especial theoretical importance because they called into question a view now widely accepted in embryology and even in child psychology, that the degree of specificity and the generality of response to stimulation is a simple function of the age or degree of maturation of the nervous system of the organism under consideration. It seemed important therefore, to attempt to discover whether or not different intensities of stimulation when applied in another receptor field would have similar or contradictory effects.

Pressure sensitivity was selected for investigation. In the current study again a typical series of guinea pig fetuses of known insemination age were used. The esthesiometers used in this study exerted pressures of known intensity and were roughly calibrated in relation to the lower threshold of each individual organism studied. Formally stated, then, the problem of the present investigation was to discover whether or not punctiform pressure stimuli of differing and known intensities had any effect on the behavior elicited by such stimulation which could be shown to be a function of the intensity of the pressure stimuli used. A large number of hair esthesiometers of the type described by von Frey (5, 6) were constructed. These were then calibrated and 17 were selected which

---

*Received in the Editorial Office on July 10, 1938.

formed a more or less continuous pressure series. In tension they ranged from .14 to 7.75. The units of *tension* here employed were determined by the use of the formula $\dfrac{\textit{milligrams of pressure}}{\textit{micra of radius of the hair}}$. Maximal force was determined for each hair by the use of a chemical balance. The radius of the hair was measured by the use of a micrometer scale in a compound microscope. Table 1 gives the tension values for the total series used. The relatively unfamiliar unit of tension is employed in these calibrations because von Frey demonstrated that for stimuli of area less than 2 sq. mm., physiological equivalence is proportional not to the area, but to the radius of the stimulus. As the dimensions of this magnitude are those of surface tension, von Frey used the term *tension* for the unit of calibration of stimulus hairs.

TABLE 1

CALIBRATION OF HAIR ESTHESIOMETERS

Number of esthesiometer in protocol	Radius in micra	Total force in milligrams	Tension in mg/μ
I	39	5.5	0.14
II	31	5	0.16
III	30	6	0.2
IV	35	10	0.29
V	38	13.5	0.35
VI	44	21	0.48
VII	52	31	0.6
VIII	49	39	0.8
IX	51	53	1.0
X	46	64	1.4
XI	43	71	1.65
XII	53	117	2.2
XIII	60	175	2.9
XIV	59	210	3.6
XV	60	275	4.6
XVI	69	414	6.0
XVII	80	620	7.75

The fetal organisms were prepared for study by a technique previously described (2, 3) by means of which placental circulation is maintained. The apparatus used, including an electrically driven moving picture camera, is shown in Figure 1. The environmental conditions of each fetus were also held as constant as possible by

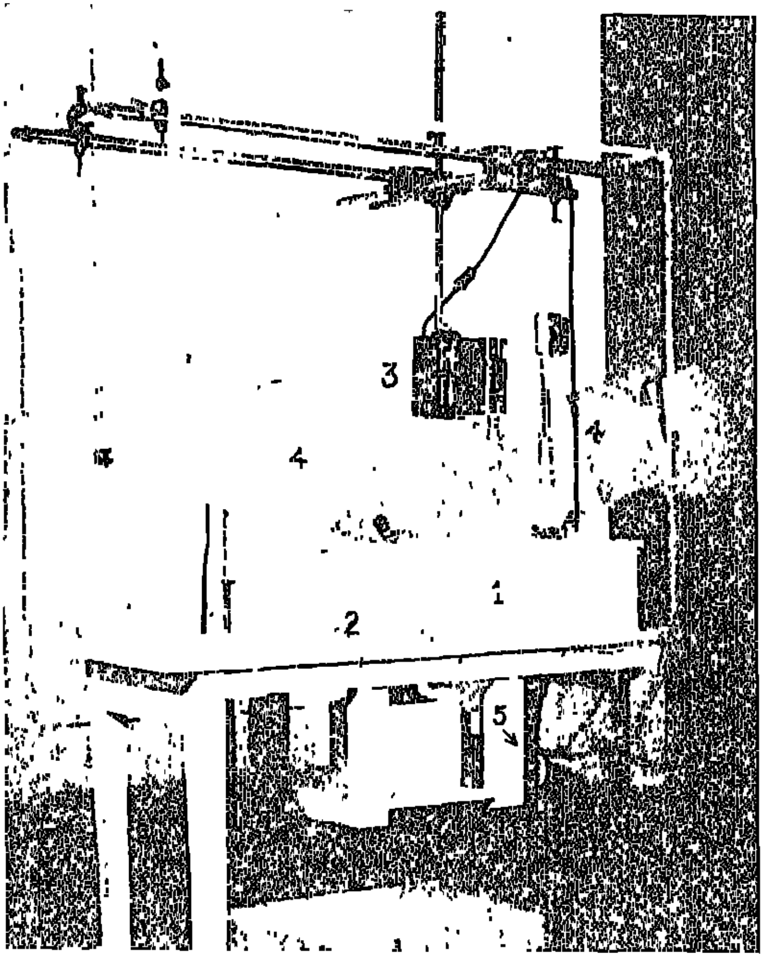


FIGURE 1

Photograph of Apparatus

1 Constant temperature tank in which is inserted tray filled with physiological saline solution Maternal animal is held in this bath by suitable supports and fetal organisms are then exposed for study

2 Holder and graded series of esthesiometers as explained in text

3. Electrically driven motion picture camera for recording

4 Reflectors for two 1000 watt lamps used to illuminate the field

5 Foot operated switch to start and stop the recording motion picture camera

keeping them immersed or almost immersed in a constant temperature bath of physiological saline solution

A series of points on the body surface were selected on the basis of previous study as typical reflexogenous zones, or points to be used in the comparative study The points selected for the present study are shown in Figure 2. In verbal description they were

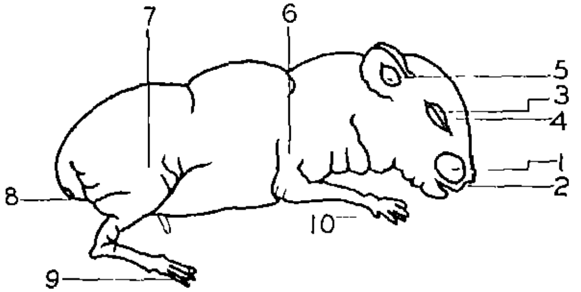


FIGURE 2

(1) The mid-point of the vibrissae pad on the snout.
(2) The inner surface of the vestibule of one external naris (nostril)
(3) The mid-point of the upper eyelid
(4) The mid-point of the lower eyelid
(5) The mid-point of the concha of the ear
(6) Skin over the shoulder joint of the pectoral girdle
(7) Skin over the hip-bone joint of the pelvic girdle
(8) Anus
(9) Point in medial plantar surface of one hind paw
(10) Point in medial palmar surface of one fore paw

Previous experiments had indicated that the vibrissae pad was one of the earliest and most sensitive cutaneous reflexogenous zones of the fetal guinea pig This well-marked region was therefore selected as the one on which to roughly calibrate the esthesiometers to be used In general, this calibration consisted in beginning with a very light stimulus hair and progressing by single steps to esthesiometers exerting greater and greater tension, until one was found which just elicited a noticeable behavioral response This roughly determined just-supraliminal esthesiometer, or in a few cases, one just above it, was then used throughout the experiment on that particular fetus and called the "light stimulus." A "heavy stimulus" was then selected. This stimulus was usually selected arbitrarily from 7 to 9 points higher in our scale of esthesiometers than the light stimulus which was selected as described above. The authors are well aware of the inaccuracy of a threshold determined in this way, but because of special conditions such as sensory fatigue and the transient optimal period for fetal study, it seemed the best method available.

The technique of stimulation consisted in applying the light or heavy esthesiometer to the spots indicated. In half the cases, the heavy stimulus was used first. In the other half the light stimulus was first applied. In so far as possible, the pressure of the esthesiometer was given by using the standard method recommended by von Frey. The resulting responses of the fetus were observed and agreed upon by two observers. A description of the responses was then recorded in short-hand. Moving picture records were also taken of the behavior released by each stimulus. As noted above, a foot-controlled electrically driven moving picture camera set just above the experimental bath was used for these recordings. This procedure made it possible for the experimenter to stimulate the animal and manipulate the camera at the same time. In each case, the same spot was stimulated three times by the strong, and three times by the light esthesiometer. In a few instances the whole series of spots was stimulated by one and then by the other of the esthesiometers, but in most cases, the two were used alternately in stimulating first spot "one" and then spot "two" and so on.

In the experiment here reported, it was especially necessary to have fetuses in absolutely optimal condition, and so in the final calculations, only 45 fetuses were used, although more than twice as many were prepared for study. For convenience of presentation of results, the organisms studied are arranged in six age groups as shown in Table 2.

No one can deny that it is difficult to quantify the responses made by a total intact organism. In this study such quantification was undertaken by dealing with the dictated protocols checked by motion pictures. Following this plan, individual items of behavior are counted. Each act of behavior which was to be considered as an independent item was decided upon in advance of the tabulation. Figure 3 shows the total number of responses in each of the six

TABLE 2
FETAL LITTERS

Lab. No. of g. pig	No. of fetuses	Av. length of litter in mm.	Av. weight litter in g.	Post insemination days
Age Group I, 35-36 days				
126	1	39.4	4.8	35
122	2	41.5	5.5	35
31	2	41.2	5.5	36
124	3	41.7	5.2	36
132	3	40.5	5.3	36
Age Group II, 41-42 days				
98	2	55.3	11.7	41
24	3	55.6	11.4	41
88	1	56.4	14.3	41
84	3	56.7	15.0	41
29	2	57.7	14.8	42
Age Group III, 45-47 days				
104	2	68.7	21.8	45
27	3	70.2	25.1	45
68	3	71.7	30.3	47
101	3	74.2	34.2	47
Age Group IV, 50-51 days				
7	2	81.3	36.2	50
34	2	70.2	29.8	50
75	3	72.4	33.2	51
92	2	72.7	35.7	51
Age Group V, 55-56 days				
56	3	72.5	31.5	55
74	2	81.5	48.4	56
130	3	80.1	41.3	56
44	1	90.7	66.9	56
70	2	92.3	62.8	56
Age Group VI, 61-62 days				
61	2	94.1	61.4	61
25	2	94.7	64.7	61
773	1	99.5	91.8	61
114	2	99.8	70.0	61

typical fetal age groups studied. This diagram also shows the percentage of responses released by strong, as contrasted with weak stimuli. It will be noticed that in percentage of total movements at all of the age groups, the strong stimuli released approximately one-third more responses than did the weak stimuli. The constancy of this result seemed surprising, especially in view of the fallibility of

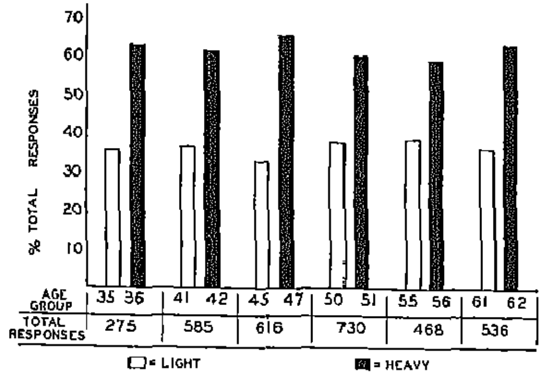


FIGURE 3

TABLE 3

PERCENTAGE OF RESPONSES AT EACH FETAL AGE-GROUP STUDIED
(Normal gestation period is 68 days)

	Group 35-36 days	
Total responses 275	Light 36.5%	Heavy 63.5%
	Group 41-42 days	
Total responses 585	Light 37.5%	Heavy 62.5%
	Group 45-47 days	
Total responses 616	Light 33.6%	Heavy 66.4%
	Group 50-51 days	
Total responses 730	Light 38.4%	Heavy 61.6%
	Group 55-56 days	
Total responses 468	Light 39.7%	Heavy 60.3%
	Group 61-62 days	
Total responses 536	Light 35.4%	Heavy 64.6%

the quantification of behavior by the verbal technique just described. Table 3 shows these graphic results in tabular form.

In summary of this quantitative data, it may therefore be said that at all fetal ages studied in this in some ways typical fetal mammal, and also in all of the typical reflexogenous zones studied, more intense pressure stimulation tends to release responses involving larger or more general patterns of response than do less intense pressure stimuli. One example of such postural difference is shown in Figure 4.

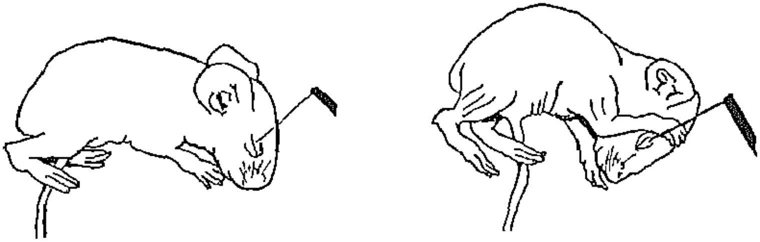


FIGURE 4

Outline Tracings from Moving-picture Film of Two Postures of 51-day Guinea-pig Fetus

Picture at left shows resting posture of the fetus as stimulated with light esthesiometer. Only response in this case was an eye-wink. Picture at right shows one phase of movement elicited by a heavier esthesiometer. In this case, the entire trunk and all limbs are involved in the reaction. Note the precise localization of the stimulated spot by the fore-paw.

These tracings are given as samples of hundreds of other postures which have been studied on the research film itself.[1]

This finding supports an observation previously made by one of the authors, namely, that in the guinea pig, specific reflexes are established early in fetal life. The present study also confirms another previous finding concerning the importance of controlling the intensity of stimulation in the study of stimulus released behavior in fetal life. To put this in another way, the present results show that internal stimulation in fetal life may mask specific reflexes and make them appear as component parts of more general patterns of behavior. In much work on fetal behavior the factor of intensity of stimulation has not been especially controlled. Hooker (7), however, has emphasized the importance of this factor.

From one point of view, it may be said that the distinction between the two intensities of stimulation in regard to specificity and generality of response, is *qualitatively* (or in an as yet unmeasured way) less clear in late fetuses, than in early ones, although quantitatively, as measured here, the relative number of responses released by weak and strong stimuli is the same in this period as in earlier periods. That is, in the late period a weak stimulus may initiate a very general adaptive response on the part of the whole organism. Such an adaptive response of the total individual could not be so elicited at an early period in fetal life This is due to the fact that unquestionably, the growth of the nervous system during active fetal life does change the character of the elicited responses There is excellent reason to suppose, by analogy with other fetal types, that as Barcroft has shown (1), oxygen tension of the blood involved in cerebral circulation changes markedly during fetal life In the late fetus, that is, much less oxygen is available than in earlier periods of motile fetal life. In the early period, neurological evidence seems to suggest that the specific reflexes are mediated more or less directly by relatively specific central nervous system connections In the late guinea pig fetus, on the other hand, cortical control and exteroceptive stimulation has in some measure at any rate, been established This leads to changes in the character of specific limb, digit, and trunk movements The terms "facilitation" and "inhibition" can be applied to some of these changes, and the total result is, if for a moment we use teleological description, that the total integrated organism seems to "attempt" to avoid the stimulus, for example. Throughout the later fetal period, the great significance of stimuli which are themselves produced by previous muscle movement must also be taken into consideration Very often, for example, a light cutaneous stimulus applied to any reflexogenous zone may lead to a local and precise reflex The extension of the digits or of a single digit, is an instance of such a response As soon as this extension has taken place, however, proprioceptive stimulation is brought about by the movement and this may itself lead to more general patterns of reaction Thus, some of the elaborate patterns of response which will be seen in the motion pictures which are to be presented must be considered not as the simple result of the stimulation of a single point on the skin, but rather as resulting from behavior released by elaborate proprioceptive as well as external stimulation Moreover, such responses are not merely patterns of *simultaneous* behavior acts, but often involve elaborate time patterns of response Such time patterns especially depend upon movement produced stimuli as well

---

[1] The writers wish to express their gratitude to the National Research Council which made the purchase of this film possible.

Certain reflexogenous zones, including mucous membrane areas such as the inner surface of the vestibule of the external nares and the anus, also seem to release somewhat more general patterns of response, no matter what the intensity of the stimulus may be These areas also seem to have a more constant lower threshold than do typical cutaneous areas. Possibly this result, which is as yet not fully established quantitatively, is due to mechanical factors such as the fact that skin thickness and hair growth vary more markedly in the cutaneous areas than in so-called mucous membrane reflexogenous zones.

In summary, it can be said that the present study employing pressure stimuli confirms the previously reported results based upon temperature stimulation concerning the importance of the factor of the *intensity of the stimulus* as a determiner of the specificity and generality of responses in fetal life. Moreover, the present results show that the same relative proportion of general and specific responses seem to be released at all of the typical fetal ages studied. This finding adds further confirmation to the view that specific responses are early developed in fetal life. Thus it becomes more than ever clear that the widely accepted formula that all specific behavior develops from the individuation of previously more general patterns of behavior needs definite revision. Such revision can come only from continued study of fetal behavior as released by stimuli of quantitatively known character.

REFERENCES

1. BARCROFT, J. The Brain and its Environment. New Haven: Yale Univ. Press, 1938. Pp. 111.
2. CARMICHAEL, L. Origin and prenatal growth of behavior. In *A Handbook of Child Psychology*, ed. by Murchison, C. Worcester: Clark Univ. Press, 1933. (Pp. 31-159.)
3. ———. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life. *Genet. Psychol. Monog.*, 1934, **16**, 338-491.
4. CARMICHAEL, L., & LEHNER, G. F. J. The development of temperature sensitivity during the fetal period. *J. Genet. Psychol.*, 1937, **50**, 217-227.
5. VON FREY, M. Beitrage zur Physiologie des Schmerzsinns. *Berichte uber die Verhandlungen der Koniglich Sachsischen Gesellschaft der Wissenschaften, Mathematisch-Physische Classe*, 1894, **46**, 185-196.
6. ———. Untersuchungen uber die Sinnesfunctionen der Menschlichen Haut. *Abhandlungen der Mathematisch-Physischen Classe der Koniglich Sachsischen Gesellschaft der Wissenschaften*, 1897, **23**, 175-266.
7. HOOKER, D. Early fetal activity in mammals. *Yale J. Biol. & Med.*, 1936, **8**, 579-602.

Laboratory of Sensory
Psychology and Physiology
Tufts College, Massachusetts

Department of Psychology
University of Rochester
Rochester, New York

The Journal of Genetic Psychology, 1939, 54, 435-454

A TEST OF GENERAL ABILITY IN FIJI*

Department of Psychology, Stanford University

CECIL W MANN

A THE PROBLEM

For many reasons, there have been but few programs of standardized tests applied to Pacific Island primitive groups living in their own environment (18), and until 1934 (7) none at all in the Colony of Fiji This report is based upon the construction and standardization of a performance test applied to 4,000 school children in Fiji An attempt was made (*a*) to produce an instrument which would measure individual differences within the racial and language groups in Fiji, and (*b*) to use the results of the test, if possible, in the making of racial comparisons.

B. THE RACIAL BACKGROUND

The Colony of Fiji consists of a group of some 250 islands—less than one hundred of which are inhabited—lying within the south

TABLE 1
VITAL STATISTICS FOR THE COLONY OF FIJI [1936 (22)]

Race	Male	Female	Total	%
Europeans°	2,263	1,765	4,028	2.03
Fijians	49,869	47,782	97,651	49.22
P E N D °°	2,325	2,249	4,574	2.31
Indians°°°	48,246	36,756	85,002	42.85
Rotumans♀	1,413	1,403	2,816	1.42
Other Pacific Islands	1,542	937	2,479	1.19
Orientals	1,525	288	1,813	91
Others	11	5	16	——
Total	107,194	91,185	198,379	

°White race, consisting almost wholly of English, Australians and New Zealanders

°°Persons of European and Native Descent, commonly referred to as half-castes in Census Reports previous to 1936

°°°East Indians

♀Rotumans are Polynesians living in a small island about 300 miles north of the Fiji Group, but belonging geographically to the Colony

---

*Accepted for publication by L M Terman of the Editorial Board, and received in the Editorial Office on July 25, 1938

tropical zone between latitude 15°*S*, and 22°*S*, and between longitudes 177°*W* and 175°*E*. The population of the Colony—the details of which are shown in Table 1—totals nearly 200,000

1 *European Group*

Of the four main groups in the Colony, the European is the smallest It totals 4,028, mainly English, Australians, and New Zealanders, engaged in administrative and professional work and in trading Fiji is a Crown Colony, administered by a Governor and Executive Council responsible to the Colonial Office in London (21, 23) The Colonial Service is made up largely of careerists and of highly selected professional men and women The trading group consists largely of the executives of sugar-growing, merchandising, and mining concerns Most of the skilled and unskilled labor and a good proportion of the clerical work is performed by Fijians and Indians. The white population of Fiji is highly selected with respect to intelligence, the proportion of the European population in the above-average group with respect to intelligence would be far higher than would be the case in the general population of America or Australia.

2 *Half-Caste Group* (*P E N D*)

The term "half-caste" in Fiji—as in most British Colonies—is used to designate a mixture of European and native race No attempt has been made in the Census of 1936 (22) to differentiate the particulars or the proportion of the race mixtures, or to include under the term the offspring of two native races, say Fijian and Tongan There are a few Chinese and Fijian half-castes in the Colony, but very rare mixtures of Fijian and Indian. The tragedy of the half caste is that he is an outcast, received by neither his mother's nor his father's race By both, he is considered to have a stigma

> The results of racial intermingling between the more widely separated races have undoubtedly not been happy in most cases In many instances, the children of mixed unions may be the offspring of inferior members of one or other race or both and so have received a poor inheritance The domestic environment of a half-caste population is, as a rule, adverse to healthy growth Such a population is, in many cases, the offspring of illicit unions, and the children grow up without any real home

> The social environment is often equally injurious both physically and morally, prejudice, hostility, and lack of opportunity may stunt and warp character (12)

3 *Fijian Group*

The Colony of Fiji is on the border line between the Polynesian and Melanesian race groups To the east are the Polynesians, tall copper-colored wire-headed people with more than a suggestion of oriental in their features; to the west are the Melanesians, shorter in stature than the Polynesians, much darker skinned, fuzzy-headed, with typical negroid features.

The history of Fiji before the Deed of Cession in 1874 is the story of the struggles of a number of small kingdoms for political supremacy, the rapid ascendency of Bau under the lead of Cakobau, and his attempt—strenuous but futile—to unite the group in a strong native kingdom (19)

The Fijians up to the middle of the Nineteenth Century were fierce warriors and cruel cannibals, their victories in war were gained by cunning or treachery and celebrated with horrible orgies of bloodshed (19) Socially they were highly developed, they maintained a system of law and order, their language was an adequate system of communication, native arts and crafts were well in evidence, in the peaceful interludes they were industrious, building canoes and houses which were famous all over the Pacific, and practising a seasonal system of agriculture

The heroic attempt of Cakobau to unify the people and to establish an island kingdom failed, and the natives became rather seriously involved financially The Council of Chiefs decided to offer the colony unconditionally to Great Britain. Two previous offers had been made to Great Britain and one to United States, but neither country was, at that time, particularly anxious to extend a Pacific Empire In 1874, by the Deed of Cession, Fiji became a Crown Colony The British Government has attempted to govern by using the old chiefly and social system of the Fijian, and in its administration of native affairs has always had the benefit of the knowledge and the wisdom of the ablest Fijian chiefs

Fiji is one of the few native areas where the coming of the white man has not been followed by depopulation of the native races The Nineteenth Century in Fiji was marked by very serious depopulation (24) but during the last 40 years there has been a steady increase in the native population due, in large measure, to the wise educational methods of the Government and the Missions with respect to the welfare of children.

It would appear that the Fijians will continue to exist as a race for some considerable time The emigration of Fijians is very seriously restricted by Government Ordinance, the birth rate over the last 50 years has been fairly constant while the death rate over the same period has declined, and there is, at present, comparatively little tendency for them to be assimilated by the alien groups in the Colony

4 *East Indian Groups*

As early as the middle of the Nineteenth Century, and even before peaceful conditions were established in Fiji, commercial groups commenced to explore the possibility of Fiji as a source of tropical produce For some time, cotton was the most important crop, to be superseded later by copra[1] and sugar-cane. The latter has been for some decades and still is the basic commercial crop It was at once obvious to the commercial interests that labor would be one of the main obstacles to commercial progress. The native Fijian viewed the coming of the trader with perfect equanimity A bountiful nature contrived to provide him with all and more than was necessary to enable him to live in comfortable circumstances with the minimum amount of manual labor Thus, he could not be coaxed or argued into wage service for an alien, and the Government fortunately forbade his coercion. For a time, natives not so fortunately placed were brought in from other Pacific Island groups.

In 1879 arrangements were completed with the Indian Government for the immigration into Fiji of large numbers of East Indians to supply the labor for the sugar-cane plantations The Indians were indentured for a period of five years with the right of repatriation at the end of the period The indenture system continued until 1916 and more than 60,000 Indians were introduced to the Colony (11) Approximately half this number were repatriated, the rest remained to bring up their families in Fiji In 1936, 72 per cent of the total Indian population were native born (22).

The term "Indian" is a loose one covering many diverse religious and language groups. These two factors, religion and language, split the Indian groups with definiteness, and form uncompromisingly rigid divisions. The majority of Indians are by religion Hindu, most of the remainder are Mohammedans. Hindi is the official language of the Indian group, but there are eight or ten vernaculars spoken by groups of varying size.

---

[1] Copra is the dried kernel of the coconut, used in making vegetable oils

The Indian population is on the increase. The birth rate is higher and the death rate lower than those of the Fijians, and at present net rate of increase it is likely that in a few years the Indian population will outnumber that of the Fijian. One of the major problems of the administration in Fiji is to preserve the letter and the spirit of the Deed of Cession wherein the welfare of the native Fijian was to be the primary concern of the Government, and, at the same time, to give full regard to the needs and rights of the industrious Indian population.

The Fijians and Indians live amicably together with no tendency towards racial mixture. The Fijian is in the enviable position of a landowner, four-fifths of the land in Fiji is secured in perpetuity by an ordinance of 1909. He has a measure of economic and social security which would be the envy of many of those in communities where rugged individualism is the basis of the economic struggle. The Indian is primarily a peasant farmer, and the great majority of Indians in Fiji till the soil. As well, during the last 20 years, the Indian has prepared himself for professional and commercial undertakings, and has proved quite successful in most of the vocations to which he has turned his attention.

C. EDUCATIONAL OPPORTUNITIES

For the European children, the Government has provided adequate educational facilities up to college entrance. The many who require additional facilities seek them in the Universities outside the Colony.

One of the first achievements of the Methodist missionaries who arrived in 1835 was the reduction of the Fijian language to writing and the publication of an English-Fijian Dictionary. Schools were soon established and for nearly 100 years the missions—mainly Methodist and Roman Catholic—have borne the brunt of the educational work of the Colony. To what purpose will be realized from the fact that 85 per cent of the natives are able to read and write their own language, and a fair percentage speak, read, and write English. The schooling situation is one that is perfectly familiar to the Fijian.

The Fijian schools vary in size and in efficiency. Ten years ago there were very few villages in Fiji without schools; the recent plan of consolidation has removed many school buildings, but there are few villages lacking educational facilities. In most of the larger centers, there are schools for some type of higher education. The Government operates a fine Medical School at Suva and a Training College at Natabua. The Missions have provided Training Colleges, Agricultural and Technical Schools, and boarding schools for girls.

For 20 years subsequent to the arrival of the first Indian immigrants there were no provisions made for what was originally conceived to be a transient group. The last 40 years, however, have seen a steady increase in the facilities for Indians and a growing realization on their part of the value of education. As yet, however, but a very small proportion of the Indian population of school age takes advantage of the facilities offered. The main problem has been to get Indian parents to allow their girls to continue at school. Up until a few years ago it would have been a rare thing to find Indian girls above the age of 10 years in schools other than the few boarding schools maintained by the Methodist Mission. The whole social and religious structure of India is against co-education, and the Government and the missions have had to proceed very cautiously. Religion, marriage customs, racial prejudice, and a conservatism of thousands of years are all obstacles in the path of the educational reformer. For many reasons, it is felt that the time is not yet ripe to put into operation the compulsory provisions of the Educational Ordinance of 1929.

The total number of children of all races attending schools in 1935 in Fiji was 20,344. (This number included only children of the statutory school age—from 6 years to 14 years.)

TABLE 2

Percentage of Total Population of School Age (6-14 Years) Actually Enrolled in the Race Groups [1935 (20, 22)]

Race	Boys	Girls	Total
European and P.E.N.D.	90.25	79.06	84.51
Fijian	72.69	56.74	64.91
Indian	37.94	13.53	26.15
Total	56.43	36.89	46.94

It will be seen from the enrollment percentages that three-quarters of the Fijian boys and little more than one-half of the Fijian girls of school age are enrolled, while little more than one-third of the Indian boys and less than one-seventh of the Indian girls are in school

The Fijian group in school probably constitutes a fair sample of Fijians There is no evidence that the Indian grouping is typical, and some evidence which indicates that it is not. Under the present conditions there is no method of getting a typical sample of the Indian juvenile population, and one must be careful in handling results It is obvious that racial comparisons made on the basis of the present selection are not much value

D. THE MENTAL TESTING PROGRAM

It has been considered necessary to outline the social and educational background of the peoples of Fiji so that the significance of the testing program may be more fully appreciated

The Fijians are primitive peoples in one sense, it is true, but they are far removed by culture-contact from the primitiveness of, say the Aborigines of Australia or the Papuans of New Guinea The racial pride of the Fijian is very strongly developed, he feels that by the Deed of Cession the welfare of his people is the most important concern of the Government, and the beneficent attitude of the Government and mission has done much to sustain this belief. His group is numerically the greatest, and, unlike so many colored groups which have contacted the white, apparently has no desire to ape the white or to live like one At the same time, the Fijian is quite progressive and is anxious to embrace any movement which will benefit his race and can be incorporated in its social structure

The Indian would justly resent the suggestion that he is a primitive He would stress the fact that while his culture is different from that of the west, it is far older He would even suggest that some of the most fundamental elements in western culture were borrowed from or had their roots in the east. The Indian has adopted rapidly and with ease his new material environment, for the majority of Indians in Fiji, India is merely the land of origins which the fortunate may visit but to which few will return as a permanent home

Neither Fijian nor Indian is to be regarded as a depressed population, in both ethnic pride is strong, both look upon the white man as an equal but of a different culture. Both races are familiar with the school situation and with graphic and pictorial representation. Few of the purely racial limitations which apply to testing programs among primitive and depressed peoples are present in Fiji (1).

The many languages in Fiji preclude the possibility of administering a verbal test, even if the translations made such a course feasible. It should be noted that there is no "pidgin" English in Fiji. The Fijians all speak Bauan Fijian, the Fijian of the island of Bau became official with the political ascendency of that kingdom and was adopted by the Government at the time of the Deed of Cession. The variation in dialects from place to place is no greater than the variation in the English spoken by Americans in Northern and Southern States. For the Indian, Hindi is the official language, but there are many who neither speak nor read this language, and, in any case, for many reasons, the standard of written Hindi in the Colony is low. For these reasons, it would seem inevitable that any testing program would have to employ non-verbal or performance material.

Agriculture is the primary industry in Fiji. Three-quarters of the Indians and almost all the Fijians are engaged in raising commercial or domestic crops. There are few towns as we know them—comprising heterogeneous groups of different families, races, and religions. Most of the Fijians live in towns (*koro*) which are communities bound by strong social and kinship ties. The Indians live in or adjacent to the large towns in separate family entities or in small patriarchal communities in the sugar-cane areas. Many of the elements of a complex civilization are absent, while there are many elements present which would not appear in a complex community.

It is certain that a test which proved satisfactory in the community on which it was standardized may not prove satisfactory in a new environment. It is most unlikely that a test standardized for American or Australian conditions will prove satisfactory in Fiji. At the same time, tests can be constructed which involve the same aspects of mental activity for members of primitive and complex communities alike. While the material over which intelligence works may be different, the quality, though not necessarily the quantity, of intelligence may be the same The kind of thinking which enables a young Australian to substitute a brass paper-fastener for a mislaid gold stud is probably the same as that which a Fijian employs when he uses a *via* leaf to keep his hair dry In both cases there is an eduction of relations

To overcome the difficulties of language and environment it was decided to construct and standardize a picture test which would use material familiar to all the groups in Fiji

1. *The Fiji Test of General Ability*

A non-verbal test was constructed using pictures and geometrical diagrams and including a test of digit substitution As a preliminary step, an extended form was constructed containing more items than were thought necessary for the final form This was administered to 750 children of all races between the ages of 10 and 14 years attending schools in and around Suva During the administering of the test, care was taken to allow every child sufficient time to complete all the tests and a record was taken of times taken by the slowest and the fastest pupils

This form was administered and scored by the writer and an examination was made of the percentage passing correctly each item. Duplicates and items that were too easy and too difficult were omitted and the remainder arranged in order of percentage difficulty No attempt was made to arrange an equal-interval scale for any of the tests employed

In the final form, the following tests were used (8)

1 Completions
2 Classifications
3 Similars
4 Opposites
5 Analogies
6 Number Series
7 Substitutions

Tests 1-5 were work limit tests, Tests 6 and 7 were time limit tests Each test, with the exception of Test 7, contained 24 items Test 7 was a digit-symbol substitution test containing 144 items

a Test 1 Completions In this test figures are arranged in serial order The problem is to discover the relationship between the figures and to complete the series

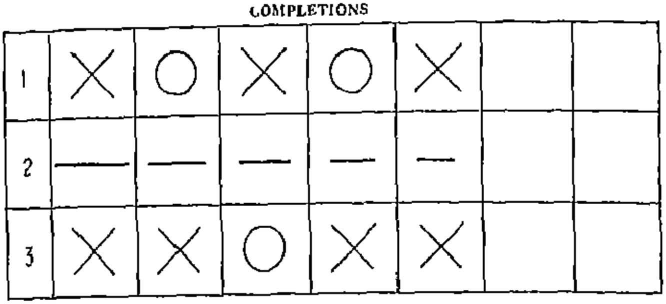


FIGURE 1
SAMPLES OF COMPLETION TEST

b *Test 2: Classifications.* In this test a number of drawings is made of objects which are common in the Fijian environment. Four of the objects are similar in some respect, the fifth is different. The problem is to discover the one that is different and to cross it out.

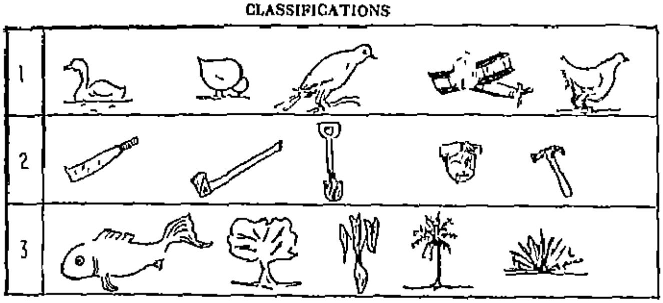


FIGURE 2
SAMPLES OF CLASSIFICATION TEST

c. *Test 3: Similars.* In this test there are five drawings of common objects, one placed alone. The problem is to discover from among the remaining four objects the one which is most like the one which stands in the column by itself, and to cross it out.

d. *Test 4. Opposites.* In this test, the problem is to cross out the object among the group of four which is most unlike the object in the same line which stands by itself.

e. *Test 5: Analogies.* In this test, the drawings are placed in

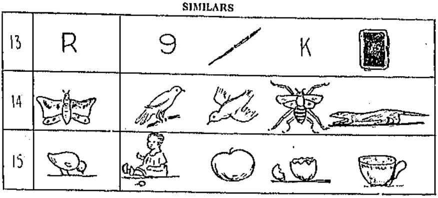


FIGURE 3
SAMPLES OF SIMILARITY TEST

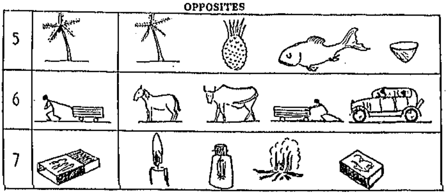


FIGURE 4
SAMPLES OF OPPOSITES TEST

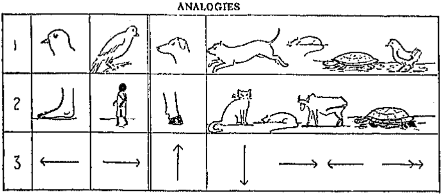


FIGURE 5
SAMPLES OF ANALOGIES TEST

four columns. The problem is to discover and to cross out the object in the group of four in the fourth column to which the object in the third column bears the same relationship as that in the first column bears to that in the second.

f. Test 6 Number series. In this test, there are 24 rows of numbers. The problem is to discover the relations which exist between the given numbers in each series, and then to educe the correlates.

NUMBER SERIES

5	6	7	8	9	10	——	——
9	10	11	12	13	14	——	——
6	8	10	12	14	16	——	——
8	8	6	6	4	4	——	——
1	7	2	7	3	7	——	——
3	6	9	12	15	18	——	——

FIGURE 6
SAMPLES OF NUMBER SERIES TEST

g. Test 7: Substitutions. In this test certain symbols are to be substituted for digits. The key is given and this is before the candidate the whole time. There are 144 substitutions to be made.

Each test is preceded by a carefully graded practice page, which

SUBSTITUTIONS

1	2	3	4	5	6	7	8	9
~	⊥	⊐	L	U	O	Λ	+	=

3	1	2	1	3	2	1	4	2	3	5	2	9	1	4
6	3	1	5	4	2	7	6	3	8	7	2	9	5	4

FIGURE 7
SAMPLES OF SUBSTITUTION TEST

has the advantage of allowing the child to become familiar with the situation and the problem. The tests are accompanied by instructions printed in English, Fijian, and Hindi (9). Since the material is non-verbal it is obvious that the instructions may be given in any language. There is no need to employ any pantomime during the test. The person giving the test should see that during the practice period the child has the opportunity of becoming thoroughly familiar with the material and the problem. The blackboard may be used during the instruction period, questions may be asked and answered, and every effort made to ensure that the child understands the situation so that he will be able to put forth maximum effort when he comes to the test proper. The practice pages preliminary to the tests proper each contain six examples, and these are to be worked over by the pupil and corrected by the teacher before the test proper is attempted.

2. *The Application of the Test*

The *"Fiji Test of General Ability"* was applied to children in 92 schools, and altogether over 4,000 children of all races were tested. The tests were administered by teachers, and great care was taken to ensure that the tests were sent to schools staffed by careful teachers who were sympathetic to the program. Schools of all sizes and in all provinces were included in the scheme and an effort was made to get as typical a sample as possible. For reasons which have already been given, this was impossible in the case of the Indian children.

TABLE 3
SHOWING NUMBERS IN RACE-AGE GROUPS TESTED

Age-group	European	Half-caste	Fijian	Indian	Total
8 yr. 6 mos.	39	—	—	—	
9 yr. 6 mos.	22	14	98	76	
10 yr. 6 mos.	30	28	222	199	
11 yr. 6 mos.	51	36	257	182	
12 yr. 6 mos.	34	39	350	183	
13 yr. 6 mos.	26	34	367	143	
14 yr. 6 mos.	33	26	303	117	
15 yr. 6 mos.	18	—	265	97	
16 yr. 6 mos.	—	—	235	93	
16 yr. 6 mos. +	—	—	390	79	
Total	253	177	2,487	1,169	4,086

The European and half-caste children were all from schools in urban centres. The Fijian and Indian children were drawn from schools of all sizes and from all possible localities in Fiji, some from the large schools in Suva and in the sugar growing centres, others from small village schools on distant islands. In every case, however, the tests were applied by a European teacher or by a native teacher who was known to the writer as a person able to follow the instructions to the letter. The teachers scored the tests and returned the raw scores. A check was made of the results of all schools which took part in the program.

It was decided to weight each test in terms of the *SD* by multiplying the scores in each test by a whole number which would make the *SD* of all tests approximately equal. For this purpose the results of the 12 yr., 6 mos., group of children were selected as being the most typical. Children of this group are found scattered throughout the whole of the elementary schools, and it is also the largest group tested.

TABLE 4
RAW SCORE RESULTS FOR 12 YR., 6 MOS., GROUP

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7
Mean	9.01	14.01	15.53	12.71	10.30	8.93	75.58
σ	2.04	4.01	5.63	4.49	2.92	2.65	19.32
Conversion factor	×2	×1	×1	×1	×2	×2	÷ 4
Total Scale score	48	24	24	24	48	48	36

As one measure of the validity of the test, the most reliable and careful teachers in four schools were asked to grade groups of children tested in order of intelligence. The coefficients of correlation between the subjective estimate and the corrected mental test scores were .68, .75, .78, .89. As another measure, 145 children of the European Grammar Schools (English-speaking) were given the Otis Self-Administering Test (Intermediate Form) and the coefficient of correlation between the Otis and the Fiji Tests was found to be 0.74± .02. To a group of 750 Fijian children the Goodenough Test (3) was applied. The coefficient of correlation between the Goodenough and the Fiji Tests was 0.78±.025.

The test was administered to the same pupils in two schools after an interval of three months had elapsed. The coefficient of correlation found between the results of the test and the re-test were (*a*) 0 89± 025 and (*b*) 0 78± 031 By the split-halves method, applied to the results from four schools the reliability was 0 87.[2]

These results seem to indicate that the test is a valid and reliable measure of general ability

3 *Results of the Tests*

Table 5 gives the scaled score results for race-age groups

TABLE 5
SCALED SCORE RESULTS FOR RACE-AGE GROUPS ON THE FIJI TEST OF GENERAL ABILITY

Age	Europeans	Half-caste	Indian	Fijian
10 6	167	137	97	87
11 6	179	144	114	103
12 6	188	153	129	117
13 6	188	153	145	125
14 6	188	162	146	132
15 6	195	—	149	138

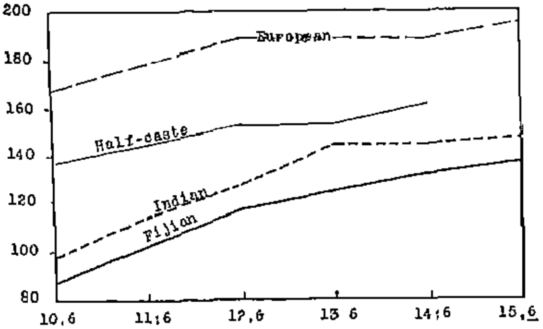


FIGURE 8
GRAPH OF RACE-AGE GROUPS USING SCALED SCORES

It will be seen that the D/σ_D values indicate that the differences in the mean scores of all the groups above except the 13 yr, 6 mos., Indian v Half-caste group are statistically significant

It will be seen that the differences in the mean scores for the four

---

[2] Using Spearman-Brown prophecy formula

TABLE 6

$\frac{D}{\sigma_D}$ FOR RACE-AGE GROUPS BASED ON A σ OF 36.0

Groups compared	12 yr 6 mos	13 yr 6 mos	14 yr 6 mos
Fijian v Indian	3.25	5.72	3.69
Indian v Half-caste	6.92	1.39	4.01
Half-caste v European	4.18	3.72	7.11

TABLE 7

VALUES OF $\frac{D}{\sigma_D}$ FOR INDIAN AND FIJIAN AGE-GROUPS

Age groups compared	Fijian	Indian	European
10.6 v 11.6	4.8	4.7	1.5
11.6 v 12.6	4.8	4.0	1.2
12.6 v 13.6	3.0	4.1	0.0
13.6 v 14.6	2.5	0.0	0.0

age groups, 10.6, 11.6, 12.6, and 13.6, are statistically significant for the Fijian and Indian races. For the European race, this is not so; it is probable that the test is more suitable for the younger European age groups than for the older.

E. INTERPRETATIONS AND CONCLUSIONS

On account of the small number of Europeans and Half-castes tested, it is not valid to make assumptions as to the value of the test when applied to these groups. For the same reason it would be unwise to make racial comparisons based on such a small number. From the results obtained, there appears to be no significance between the differences in the mean scores for the older groups of Europeans and Half-castes, although there are significant differences when the groups are compared age-for-age. It has already been pointed out that the European population in Fiji is not a population typical of America or Australia. It consists mainly of administrators and commercial executives, a group with a higher proportion of intellectually able people than one would find in the commonality. Since a majority of both groups was tested—European and half-caste children—and in view of the statistical significance of the differences in the means, one can say that the European children in Fiji rate higher on the Fiji Test of General Ability than do the half-castes

The results, however, give no warranty for a blanket statement relative to the ability of whites-in-general and the half-castes in Fiji

The results for Fijians and Indians indicate, in the first place, that there is a significance between the differences of the means for age-groups 10 6 through 13 6 It would appear that the Fiji Test of General Ability is a valid and reliable measure of ability within the race groups

In the second place, one must use caution in making racial comparisons From Table 6, it is seen that the differences in the means of the age groups 12, 13, and 14, Fijian and Indian, are significant, but one must turn back to the social and educational situation to get the real racial picture For reasons discussed previously, it is obvious that while there was measured a typical sample of Fijian children, the sample of Indian children measured is highly unlikely to be thoroughly representative. There are three times as many Fijian as Indian children in school, and three times as many Indian boys as girls in school It is likely that the Indian children that are at Indian schools are the children of enlightened parents who realize the advantages of education and take pains and make sacrifices to embrace the educational facilities offered The children who are likely to be missing are the children of peasant farmers. With a more representative samplings, the results for Indians might not be significantly higher than those of the Fijians

The graph showing the typical amount of overlap indicates that on the basis of the very faulty sampling the mean of the Indians on the 12 yr , 6 mos , group is higher than that of the Fijians, there is a substantial number of Fijians who are the equal of many Indians It would appear that while the Fiji Test is a useful measure of individual differences within the racial groups it cannot, in view of the inevitable difficulties of sampling, be regarded as a reliable measure of racial differences

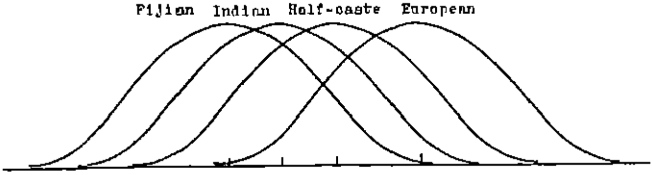


FIGURE 9

SCHEMATIC GRAPH SHOWING TYPICAL AMOUNT OF OVERLAP IN 12 YEAR, 6 MONTH, RACE GROUPS

The educational implication is not hard to find Educational problems should be solved by educational methods There is always a tendency for some to attempt to solve educational problems in diverse race groups by the application of racial methods Thus, we find in some places the philosophy that there should be racial schools each with occupations suitable for different races. The results of this test indicate that there are children of all four races in Fiji who are equal in ability, age for age, and the implication is that every child in Fiji, no matter what his color or race, should have equality of educational opportunity It may be convenient for some years in Fiji to segregate the races, but this should not be an obstacle in the way of equality of educational opportunity.

F. SUMMARY

1 An attempt was made to construct and standardize an instrument for the measurement of general ability of the Fijian and Indian school population of Fiji

2 Because of culture contact of the Fijians with whites, and because of the ancient culture of the Indians—most of whom are of the Hindu or Mohammedan religions—neither race can be regarded as primitive The majority of the Fijians and a smaller percentage of the Indians are familiar with the schooling situation and with testing conditons.

3 There is one language common to the Fijians but eight or ten different languages are vernacular to Indian groups of varying sizes The majority of both races are engaged in commercial or domestic agriculture and the material environment of the Colony is very different from that of a more complex commercial and industrial environent

4. In view of the language differences and the differences in environment it was considered necessary to construct and standardize a non-verbal test using geometrical figures and pictures of objects familiar to the children of both races

5 The Fiji Test of General Ability was administered to 2,487 Fijians and to 1,169 Indians From statistical considerations it would appear to be a valid and reliable measure for the measurement of the general ability of the two races

6. The circumstances surrounding the sampling of the Indian children were such that no reliable racial comparisons can be drawn

from the results of the test Of the Fijian children of school age nearly 65 per cent are in school and a representative sample of these was tested Only 26 per cent of the Indian children of school age are in school. The mean scores for the Indian age-groups are higher than those of corresponding Fijian age-groups, but racial comparisons are invalidated by uncontrollable selective factors.

REFERENCES

1 ANASTASI, A. Differential Psychology New York, Macmillan, 1937 Pp 615.

2 BURTON, J W, & DEANE, W A Hundred Years in Fiji London Epworth Press, 1936 Pp 144

3 GOODENOUGH, F Measurement of Intelligence by Drawings New York World Book, 1926 Pp 177

4 KLINEBERG, O Race Differences New York Harper, 1935 Pp 367.

5 LEITER, R G The Leiter international performance scale (Appendix by S D Porteus) *Univ Hawaii Bull*, 1936, 7, 1-42

6 LIVESAY, T M Racial comparisons in performance on the American Council psychological examination *J Educ Psychol*, 1936, 27, 631-634

7 MANN, C W Education in Fiji Australian Council Educ Res Ser No 33 London Oxford Univ Press, 1935 Pp 138

8 ——— Fiji Test of General Ability. Suva Government Printer, 1935 Pp 24

9 ——— Handbook of Instructions, Fiji Test of General Ability Suva Government Printer, 1935 Pp 22

10 ——— Objective Tests in Fiji Suva Government Printer, 1937 Pp 40

11 MCGUSTY, V W T East Indians in Fiji Annual Medical and Health Report Fiji Government Printer, 1932 (Pp 44-46)

12 OLDHAM, J Christianity and the Race Problem London Epworth Press, 1930 Pp 186

13 OLIVER, R A C Adaptation of intelligence tests to tropical Africa I & II *Oversea Educ*, 1933, 4, 186-191, 5, 8-13

14 ——— Comparison of abilities of races, with special reference to East Africa *East African Med J*, 1932, 160-204

15. ——— General Intelligence Test for Africans Nairobi, Kenya Colony Government Printer, 1932

16 ——— Mental tests in the study of the African *Africa*, 1934, 7, 40-46

17 PORTEUS, S D Primitive Intelligence and Environment New York Macmillan, 1937 Pp 325

18. ——— Psychology of a Primitive People. New York, Longmans, Green, 1931

19 WILLIAMS, T Fiji and the Fijians (3rd ed) London Hodder & Stoughton, 1870 Pp 224

Official Publications of the Government of Fiji.

20 Annual Report of the Director of Education Suva Government Printer, 1936

21. Blue Book of Fiji Suva Government Printer, 1935

22 Census Report. Suva. Government Printer, 1936

23 Handbook of Fiji Suva Government Printer, 1937

24. Report of the Commission on the Decrease of the Native Population Suva Government Printer, 1896 Pp 160

Department of Psychology
University of Denver
Denver, Colorado

SHORT ARTICLES AND NOTES

The Journal of Genetic Psychology, 1939, 54, 455-456

PSYCHOMETRIC TEST RESULTS IN TWO CASES OF PRECOCIOUS PUBERTY*

New York Hospital and Cornell University Medical College

Arthur L. Benton and Frank A. Hagmann

Reviews of the literature concerning the mental status of children showing pubertas praecox have been published by Kulmann and Stone (2) and Keene and Stone (1). The reviews indicate that there are surprisingly few cases on record in which standardized psychometric tests were given to determine the actual level of intellectual function. Keene and Stone found a total of only 20 cases in which the children were actually tested. Of these 20 cases, 10 per cent were rated as above the average level, 45 per cent were rated as average, and 45 per cent were rated as below the average level in intellectual development.

In view of the paucity of accurate data concerning the intellectual development of these children, the writers believe that it may be of some interest to record the psychometric findings in two cases of pubertas praecox who were seen in the Children's Clinic of the New York Hospital and who were referred for psychometric examination.

The first child was a girl who at the age of 2 years, 10 months, was admitted to the Children's Clinic for observation. After discharge from the clinic the patient was followed in the Out-Patient Department. Socio-economic status of the family was below the average level. Somatic signs of sexual precocity were first noted at the age of 7½ months, the breasts were overdeveloped and a small amount of pubic hair was present. The child began to menstruate at the age of 2½ years and continued to menstruate irregularly thereafter. She was seen for psychometric examination when she was

---

*Received in the Editorial Office on June 3, 1938.

4 years, 5 months, old. On the Stanford-Binet (original form), she made a Mental Age score of 4 years, 10 months, Intelligence Quotient—109. On a short battery of performance tests, she made a median score of 4 years, 6 months. She was judged to be of average or high average intelligence

The second child was a girl of 9 years, 9 months, who was admitted to the Children's Clinic for observation. She was in Grade 4*B* in school, did well in her studies, and was considered a normally intelligent child Socio-economic status of the family was below the average level Somatic signs of sexual precocity first appeared at 3 months of age, at which time the breasts were overdeveloped The child began to menstruate at the age of 7 months and continued to menstruate about every 2 to 3 months thereafter. She made a Mental Age score of 9 years, 8 months (Intelligence Quotient—99), on Form *M* of the Revised Stanford-Binet

Both of these children showed at least an average degree of intellectual development

REFERENCES

1. KEENE, C M, & STONE, C P Mental status as related to puberty praecox. *Psychol Bull*, 1937, **34**, 123-133
2. KULMANN, L, & STONE, C. P Notes on the mental development of children exhibiting the somatic signs of puberty praecox *J Abn & Soc Psychol.*, 1927, **22**, 291-324.

Cornell University Medical College
1300 York Avenue
New York City

The Journal of Genetic Psychology, 1939, 54, 457-460

WORRIES OF COLLEGE FRESHMEN*

Department of Psychology, University of Rochester

RUTH LUNGER AND JAMES D. PAGE

The first year of college life is frequently a period of re-orientation and adjustment, often resulting in considerable mental stress An important contributing factor to this state of tension is worry—worry over finances, grades, personal shortcomings, choice of vocation, social success, and familial difficulties.

We have attempted to analyze some of these worries by constructing a Worry Inventory of 78 items and administering it to 100 college freshmen of each sex In addition we have investigated the relation of worries to intelligence, self-evaluation as to general superiority and personality adjustment, by giving the American Council on Education Intelligence Test and a Superiority-Inferiority Self Rating Scale to both groups, and the Bell Adjustment Inventory to the female group

Each item in the Worry Inventory was answerable in one of three ways item worried about "very much," "some," or "not at all" As indicated in Table 1, the average college freshmen, when confronted with a list of worries, admits being very much worried about five or six and is somewhat worried about an additional 22 or 25, making a total of some 30 items worried about to some degree About five per cent of the men and of the women reported more than 50 worries and an equal percentage fewer than 12 As is apparent from Table 1, sex differences were entirely negligible with respect to either incidence or intensity of worries

From one-tenth to one-fifth of each sex expressed some concern with respect to such things as developing some disease or being injured, familial difficulties, and social success Approximately one-fifth to one-fourth were worried about. social obligations, mental inferiority, being too nervous, inability to get along with other people, what happens after death, and insufficient funds. Some personal defects, the possibility of losing one's friends, not being

---

*Accepted for publication by Leonard Carmichael of the Editorial Board, and received in the Editorial Office on June 23, 1938

as happy as one felt one should be, provided cause for worry for about a third of the students Roughly about one-half expressed some concern over such items as· general religious problems, physical defects, being late for appointments, familial obligations, inability to make friends, and vocational success.

A significantly higher proportion of women than men (critical ratio being 3 or over) were worried about.

1 Not being popular socially (*W* 71%, *M* 49%)

2 The possibility that no one cares for them (*W* 32%, *M* 14%)

A significantly higher proportion of men than women were worried about:

1 Future financial security (*W* 49%, *M* 80%)

2. The possibility that they will have to support their parents in their old age (*W* 4%, *M* 18%).

3 Not being able to support those who are or those who might become dependent upon them (*W* 16%, *M* 35%)

4 Being disinterested in their work (*W* 24%, *M* 52%)

5. Being underweight (*W* 7%, *M* 27%)

6 Becoming ill (*W* 7%, *M* 30%)

In other words women were more concerned with receiving attention and affection while men were more concerned with being strong and healthy and meeting their financial obligations.

TABLE 1

Extent of Worries Reported by 100 College Freshmen of Each Sex

	Incidence				Critical
	Number reported		Per cent of total		sex ratio
Degree of worry	Men	Women	Men	Women	D/σ_{diff}
Very much	5	6	6	8	54
Some	25	22	32	28	85
Not at all	48	50	62	64	29

The proportion of students reporting individual worries varied considerably from one item to the next, but with the one exception noted below (Table 2) each of the 78 items was a source of worry for some students.

In order to obtain some measure of the *consistency of the worries*, the Inventory was given a second time to 67 of the men after an

TABLE 2

	Per cent reporting specific worries Men	Per cent reporting specific worries Women
The highest proportion of freshmen were worried about		
1. Not being as successful in their work as they would like to be	94	89
2. Hurting other people's feelings	85	85
3. The impression they make on others	76	84
4. Not working hard enough	80	76
The smallest percentage of freshmen were worried about		
1. The possibility that they are foster children	3	0
2. Going insane	2	4
3. Dying	9	9
4. Growing old	14	6
5. Their home being too shabby to invite and entertain friends	15	4

interval of one month. The coefficients of correlation between test and retest was +.76±.04 which incidentally is practically identical with the reliability coefficient of +.75 reported by McGeoch and Whitely (2) for the worry section of the Pressey *X-O* test after an interval of 45 days. On our retest there was a tendency for the subjects to report fewer worries. The group averaged three items worried about "very much" and an additional 19 worried about "some." McGeoch and Whitely found no change in the number of worries after a 45-day interval but a marked decrease after an interval of 90 days.

The various intercorrelations made are summarized in Table 3. In each instance the number of cases approximated 100.

TABLE 3

THE RELATION OF WORRIES TO INTELLIGENCE AND OTHER FACTORS

Tests	Coefficients of Correlation Men	Coefficients of Correlation Women
Worry Inventory and Superiority Self-Ratings	−.17±.07	−.36±.06
Worry Inventory and *ACE* Intelligence Test	+.11±.07	−.25±.06
Bell Inventory and *ACE* Intelligence Test		−.15±.07
Worry Inventory and Bell Inventory		+.39±.06

In general agreement with similar studies (1, 3) no relationship was found between the *ACE* Intelligence Test scores and either the Worry or Bell Inventory There was a tendency for worriers to rate themselves lower in superiority and to make higher maladjustment scores on the Bell Adjustment Inventory than non-worriers.

SUMMARY

A Worry Inventory was administered to college freshmen and the incidence of worries related to such factors as sex, intelligence, personal adjustment, and self-ratings with respect to superiority-inferiority

It was found that both men and women reported about the same number of worries, but there were certain qualitative differences. Worriers tended to rate themselves lower on a Superiority Scale than non-worriers and reported higher maladjustment scores on the Bell Adjustment Inventory The correlation between intelligence and incidence of worries was negligible

REFERENCES

1. CHAMBERS, O R Character trait tests and the prognosis of college achievement. *J Abn. & Soc Psychol*, 1925, **20**, 303-311
2. McGEOCH, J A, & WHITELY, P L The reliability of the Pressey *X-O* Tests for investigating the emotions *Ped Sem & J Genet Psychol*, 1927, **34**, 255-270.
3. THURSTONE, L L, & THURSTONE, T. G A neurotic inventory *J Soc Psychol*, 1930, **1**, 3-30

Department of Psychology
University of Rochester
Rochester, New York

The Journal of Genetic Psychology, 1939, 54, 461-465

CHANGES IN VARIABILITY IN "DIGIT-SYMBOL SUBSTITUTION" PERFORMANCE MEASURED AT THE BEGINNING AND AT THE END OF PRACTICE*

William Woods College

D G RYANS

A THE PROBLEM

Individual differences in knowledges and skills are so evident and so widely recognized that a pronouncement of their existence, even as an introductory note, is strongly suggestive of triteness Also, that the effects of specific training are influential in producing and accentuating such differences is readily observable in the data of numerous experiments and in the results of instruction. Rising out of these psychological axioms is the oft-discussed question, first called to attention by Thorndike, of variations in relative accomplishments measured at the beginning and at the end of practice presumably equal for all the individuals engaged therein Stated in other words the problem concerns itself with the inquiry do originally determined differences in a group tend to remain constant, to decrease, or to increase with equal opportunities for learning?

The approaches attempted and the results obtained by investigators in studying the effects of training upon individual differences have been many and varied Perhaps the most common method employed has been that of calculating the dispersion, in terms of the standard deviation of the mean, of a group of measures on the first trial and the last trial of a practice series Other things being equal, the discovery of a larger standard deviation at the end of practice would support the contention that differences between individuals increase with training, the finding of a smaller standard deviation at the conclusion of a learning series would be acceptable as evidence favoring the contrary view, that differences between individuals decrease with training

Actually the solution of the problem has not been so simple as it

---

*Received in the Editorial Office on June 24, 1938

may appear from the foregoing. Questions both of scoring and of method have arisen which have complicated and confused available results. For example, Chapman (3), Stoddard (10), and others point out that, in the absence of adequate experimental controls such as scales with determined absolute zero points and known intervals, no valid comparisons of the type sought are possible. Reed (8), too, argues that by chance alone the measures of dispersion of a group will increase as the scores increase, and, therefore, that standard deviations determined at the beginning and at the end of practice are not directly comparable. Rather, he believes, a measure of relative variability such as SD/M should be employed. In light of such comments, whether we accept them or not, it is necessary to note that reported results are not equally meaningful. Kincaid (6) reviews and summarizes 24 sets of data relating to the effects of training on existing differences. When the standard deviation alone is considered, individual differences appear to increase with practice in 19 of the 24 cases. When, however, relative variability (SD/M) is determined from the same raw scores, differences seem to increase from the beginning to the end of practice in 11 cases and to decrease in the remaining 13. Aside from the matter of statistics the findings of these studies may be ambiguous in light of the arbitrary nature of the scales employed for scoring.

At least two studies have been published in which the authors made use of scaling techniques before interpreting their data. Anastasi (2) found that the variability of a group of 1,000 college students definitely increased from the beginning to the end of practice at such tasks as "cancellation," "finding hidden words," "symbol-digit substitution," and "vocabulary knowledge" when the standard deviations of *both raw and scaled scores* were considered. Ewert (4), likewise, reported that when the units of measurement in "mirror-tracing" were carefully weighted for difficulty (by determining the average number of individual units per trial necessary to progress one such average unit on the total scale) individual differences tended to increase with practice.

Reed's contention, earlier referred to, that only measures of relative variability are suitable for making comparisons of a group at different stages of learning, has been challenged by Anastasi (1, 2), who holds that *absolute* measures are far more adequate and desirable. In the first place, Anastasi maintains, the standard deviation will *not*, in all cases, increase as the scores and their Mean become larger Secondly, it is indicated, relative measures of variability, being ratios, assume the equality of units of measurement, whereas, actually such equality at the beginning and at the end of practice does not exist and relative measures, therefore, give a unit less weight in earlier than in later practice Measures of absolute variability are recommended by Anastasi.

B Procedure

In the studies about to be reported two groups of subjects were employed. Both were made up of unselected Junior College sophomores, members of the writer's classes in General Psychology. Sample *A*, numbering 58, was practiced for 20 trials at a digit-symbol substitution learning problem Sample *B*, consisting of 51 subjects, was practiced for 50 trials with the same materials

The individuals participating were required to learn a key (*H*-1, *F*-2, *K*-3, *Z*-4, *M*-5) and to substitute digits for corresponding letters on mimeographed test sheets Trials were one minute in length and were interspersed by 30-second rest periods

Two types of scores were employed *first,* raw scores, or the number of substitutions correctly made *per* trial, and *second,* scores for comparable periods of practice for different individuals The technique by which the latter units were obtained consisted of the division of individual learning records into equal tenths based upon the maximum attainment of each individual during practice (11). For example, if a subject in Sample *B* obtained his largest number of substitutions during the series on trial No 42, his total number of substitutions for the first 4.2 trials and his total number of substitutions for the 4 2 trials immediately preceding and including that on which he reached his limit were recorded Means and standard deviations for the groups were calculated for the first and the last trial of a series and for the initial tenth and the final tenth of learning

Measures of *absolute* variability were employed to determine the extent of individual differences at the beginning and at the end of practice Conclusions were based upon the observed statistical *significance of the differences* between these standard deviations

C Results

The results of the comparisons attempted are summarized in the

TABLE 1

VARIABILITY IN "DIGIT-SYMBOL SUBSTITUTION" PERFORMANCE MEASURED AT THE BEGINNING AND AT THE END OF PRACTICE

(Based on raw data)

		Mean		Standard deviation				
	N	First trial	Final trial	First trial	Final trial	*Diff* (*S.D.*)	*S.D. Diff*	$\frac{Diff.}{S.D.\,Diff}$
Sample *A* (20 trials)	58	48.0	86.6	8.1	14.5	6.4	1.5	4.2
Sample *B* (50 trials)	51	46.3	128.7	8.8	18.4	9.6	2.0	4.8

accompanying tables. In Table 1, the findings derived from *raw* data are shown. It may be observed that the absolute variability of both samples seemed to increase from the beginning to the end of practice and that the critical ratios were such as to suggest practical certainty regarding the direction of the differences. Similarly, as shown in Table 2, when calculations are based upon comparable tenths of individual records, standard deviations tend to increase with practice and the differences again appear to be statistically reliable.

TABLE 2

VARIABILITY IN "DIGIT-SYMBOL SUBSTITUTION" PERFORMANCE MEASURED AT THE BEGINNING AND AT THE END OF PRACTICE

(Based on comparable tenths of individual records)

		Mean		Standard deviation				
	N	First 1/10	Final 1/10	First 1/10	Final 1/10	*Diff* (*S.D.*)	*S.D. Diff*	$\frac{Diff}{S.D.\,Diff}$
Sample *A* (20 trials)	58	125.3	225.9	22.8	38.5	15.7	4.1	3.8
Sample *B* (50 trials)	51	270.8	592.0	57.0	102.7	45.7	11.6	3.9

Such results suggest that individual differences, as measured by the *standard deviation*, become more evident after than before practice at this specific task. The conclusion, while probably valid for a wide range of human activities, is here limited to the particular materials employed and samples studied.

References

1 Anastasi, Anne Differential Psychology. New York Macmillan, 1937 Pp 615.

2 ——— Practice and variability A study in psychological method *Psychol. Monog*, 1934, **45**, No 5 Pp 55.

3 Chapman, J C Statistical considerations in interpreting the effects of training on individual differences *Psychol Rev*, 1925, **32**, 224-234

4 Ewert, H The effect of practice on individual differences when studied with measurements weighted for difficulty *J Gen Psychol*, 1934, **10**, 249-285

5 Hunter, W S Learning. IV Experimental Studies of Learning Chap. 11, *Handbook of General Experimental Psychology* Worcester, Mass Clark Univ Press, 1934 (Pp 497-570)

6 Kincaid, M A study of individual differences in learning *Psychol Rev*, 1925, **32**, 34-53

7 Peterson, J, & Barlow, M C The effects of practice on individual differences *27th Yearbook, Nat Soc Stud Educ*, Part II, 1928, 211-230

8 Reed, H B The influence of training on changes in variability in achievement *Psychol Monog*, 1931, **41**, No 185 Pp 59

9 Skaggs, E B The effects of training upon individual differences Series I *J Genet Psychol*, 1936, **49**, 261-267

10 Stoddard, G D The problem of individual differences in learning *Psychol Rev*, 1928, **35**, 479-485

11 Vincent, S B The function of the vibrissae in the behavior of the white rat *Behav Monog*, 1912, **1**, No 5 Pp 81

William Woods College
Fulton, Missouri

The Journal of Genetic Psychology, 1939, 54, 467-470

AN OBSERVATION OF THE CHANGES IN VARIABILITY OF HIGH, MEDIUM, AND LOW "INTELLIGENCE" GROUPS, MEASURED AT THE BEGINNING AND AT THE END OF "DIGIT-SYMBOL SUBSTITUTION" PRACTICE*

William Woods College

D G RYANS

In a previous report (1) the writer briefly discussed the effects of training upon group variability in a digit-symbol substitution situation Employing the standard deviation it was observed that dispersion appeared to increase from the beginning to the end of a practice series and that the obtained differences were statistically significant This was true both when the scattering of scores was determined for the first and the last learning trials of a series and when it was calculated for the initial and the final tenths of individual practice

Since the relationship between "intelligence" and response to training is notably apparent, and in light of the possibility that organisms possessing varying degrees of general aptitude do not profit equally from experience as measured by gains over an initial record, the question of the *differential variability of intellectual groups during learning* is suggested Does the increase in individual differences from the beginning to the end of practice affect all groups the same, or is there reason to believe that the general trend is contributed to a greater or lesser extent by diversely selected samples?

To the end of investigating this problem as stated, two samples of subjects were each subdivided on the basis of intelligence test performance into selected groups roughly classified as "High," "Medium," and "Low" The subjects were Junior College sophomores Sample *A* practiced for 20 trials at digit-symbol substitution learning and Sample *B* practiced for 50 trials on the same material In compiling the results, individual records were divided into tenths

---

*Received in the Editorial Office on June 24, 1938.

with respect to maximum achievement (see preceding description) and the variabilities of the groups calculated for the initial and final tenths of practice.

In Table 1 the results pertaining to Sample *A* are given. Table 2 presents similar results for Sample *B*.

It is noteworthy that both samples studied show greater increases in variability, measured at the beginning and at the end of digit-symbol substitution practice, for the "Medium intelligence" groups than for either the "High" or "Low." While none of the increases appeared to be statistically significant (whereas they *were* before *selection was introduced) the reliability of the differences of the* extreme groups was measurably less than that of those embracing the middle ranges of "intelligence." If our data be valid it would seem that individual differences, as measured, increase much more with training among individuals of average intellectual ability than they do among subjects classified as relatively superior or relatively inferior in "intelligence." This conclusion is limited, of course, to the samples and materials studied.

REFERENCE

1. RYANS, D. G. Changes in variability in "digit-symbol substitution" performance measured at the beginning and at the end of practice. *J. Genet. Psychol.*, 1939, 54, 461-465.

William Woods College
Fulton, Missouri

The Journal of Genetic Psychology, 1939, **54**, 471-488.

BOOKS

The *Journal of Genetic Psychology*, the *Journal of General Psychology* and the *Journal of Social Psychology*, will buy competent reviews at not less than $2 per printed page and not more than $3 per printed page.

Conditions. Only those books that are listed below in this section are eligible for such reviews. In general, any book so listed contains one or more of the following traits: (*a*) Makes an important theoretical contribution; (*b*) consists largely of original experimental research; (*c*) has a creative or revolutionary influence in some special field or the entire field of psychology; (*d*) presents important techniques.

The books are listed approximately in order of receipt, and cover a period of not more than three years. A reviewer must possess the Ph.D. degree or its equal in training and experience.

Procedure. If among the books listed below there is one that seems important to you, you are invited to write a review of that book. It is not necessary to make arrangements with the Editor. It does not matter if the book has already been reviewed.

(1937)

ALLPORT, G. W. Personality: A Psychological Interpretation. New York: Holt, 1937. Pp. 588.

BECK, S. J. Introduction to the Rorschach Method. *Res. Monog., Amer. Orthopsychiat. Asso.*, 1937, No. 1.

NEWMAN, H. H., FREEMAN, F. N., & HOLZINGER, K. J. Twins: A Study of Heredity and Environment. Chicago: Univ. Chicago Press, 1937. Pp. 369.

MURPHY, G., MURPHY, L. B., & NEWCOMB, T. M. Experimental Social Psychology. New York: Harper, 1937. Pp. 1121.

SOUTHALL, J. P. C. Introduction to Physiological Optics. New York and London: Oxford Univ. Press, 1937. Pp. 426.

TERMAN, L. M., & MERRILL, M. Measuring Intelligence. Boston: Houghton Mifflin, 1937. Pp. 461.

THOMAS, W. I. Primitive Behavior: An Introduction to the Social Sciences. New York: McGraw-Hill, 1937. Pp. 847.

STAGNER, R. Psychology of Personality. New York: McGraw-Hill, 1937. Pp. 465.

Murphy, L. B. Social Behavior and Child Personality. New York: Columbia Univ. Press, 1937. Pp. 344.

Anastasi, A. Differential Psychology. New York: Macmillan, 1937. Pp. 615.

Burrow, T. The Biology of Human Conflict. New York: Macmillan, 1937. Pp. 435.

Gesell, A., & Ilg, F. L. Feeding Behavior of Infants. Philadelphia: Lippincott, 1937. Pp. 210.

(1938)

Thurstone, L. L. Primary Mental Abilities. *Psychomet. Monog.*, 1938, No. 1. Pp. 121.

Spearman, C. Psychology Down the Ages. (2 vols.) New York: Macmillan, 1938. Pp. 454 & 355.

Gesell, A., & Thompson, H. The Psychology of Early Growth. New York: Macmillan, 1938. Pp. 290.

Strong, E. K. Psychological Aspects of Business. New York: McGraw-Hill, 1938. Pp. 629.

Boas, F. The Mind of Primitive Man. (Rev. ed.) New York: Macmillan, 1938. Pp. 285.

Sherman, M. Mental Conflicts and Personality. New York: Longmans, 1938. Pp. 319.

Stevens, S. S., & Davis, H. Hearing: Its Psychology and Physiology. New York: Wiley, 1938. Pp. 489.

Guthrie, E. R. The Psychology of Human Conflict. New York: Harper, 1938. Pp. 408.

Murphy, G., & Likert, R. Public Opinion and the Individual. New York: Harper, 1938. Pp. 316.

Katz, D., & Schanck, R. L. Social Psychology. New York: Wiley, 1938. Pp. 700.

Ellis, W. D. A Source Book of Gestalt Psychology. New York: Harcourt, Brace, 1938. Pp. 403.

Woodworth, R. S. Experimental Psychology. New York: Holt, 1938. Pp. 889.

Skinner, B. F. The Behavior of Organisms: An Experimental Analysis. New York: Appleton-Century, 1938. Pp. 457.

Seashore, C. E. Psychology of Music. New York: McGraw-Hill, 1938. Pp. 408.

Terman, L. M. Psychological Factors in Marital Happiness. New York: McGraw-Hill, 1938. Pp. 474.

Murray, H. A., *et al.* Explorations in Personality. New York: Oxford Univ. Press, 1938. Pp. 761.

Landis, C., & Page, J. D. Modern Society and Mental Disease. New York: Farrar & Rinehart, 1938. Pp. 190.

Munn, N. L. Psychological Development. New York: Houghton-Mifflin, 1938. Pp. 582.

Rashevsky, N. Mathematical Biophysics. Chicago: Univ. Chicago Press, 1938. Pp. 340.

Köhler, W. The Place of Value in a World of Facts. New York: Liveright, 1938. Pp. 418.

(1939)

ROSETT, J. The Mechanism of Thought, Imagery, and Hallucination. New York: Columbia Univ. Press, 1939. Pp. 289.

PRATT, C. C. The Logic of Modern Psychology. New York: Macmillan, 1939. Pp. 185.

DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., SEARS, R. R., *et al.* Frustration and Aggression. New Haven: Yale Univ. Press, 1939. Pp. 209.

GESELL, A., CASTNER, B. M., THOMPSON, H., & AMATRUDA, C. S. Biographies of Child Development. New York: Harper, 1939. Pp. 328.

CRITICAL REVIEWS OF RECENT BOOKS

(*Skinner, B. F. The Behavior of Organisms. New York: Appleton-Century, 1938. Pp. 457.*)

REVIEWED BY ERNST WOLF

The title of Skinner's book, *The Behavior of Organisms,* promises by its generality something in which every experimental psychologist and biologist should be greatly interested. If we understand as behavior the functional activities of an organism in relation to its surroundings, it will be possible to collect quantitative data on experimentally recognizable expressions of the activities of an organism from which a concept of behavior can be obtained. Any modifications of factors having an influence upon an animal under observation is a stimulus which will be reacted to by a readjustment of the organism to the new environmental conditions created (response). For the observable relations between these two recognizable magnitudes Skinner uses the term Reflex. The static and dynamic laws of such "reflexes" and their rôle in building up complex behavior is discussed. By behavior is then meant

> . . . simply the movement of an organism or of its parts in a frame of reference provided by the organism itself or by various external objects or fields of force. It is convenient to speak of this as the action of the organism upon the outside world, and it is often desirable to deal with an effect rather than with the movement itself.

For an analysis of behavior in these terms Skinner's numerous experiments on the albino rat are presented. The principle of his experimental procedure consists of using a response of the rat and its modifications. The behavior consists in lifting up the fore part of its body, pressing and releasing a lever, reaching into a tray, seizing a pellet of food which dropped into the tray by the action of the lever, withdrawing from the tray, and eating the pellet. The whole performance represents a pretty complex act. It is a reflex chain which for analytical purposes can be dissolved into its component parts Since, however, the number of efforts made to obtain a pellet and the velocity of repetitions of pressing the lever can be easily recorded, the numerical values of the lever actions serve as a measure which can conveniently be treated analytically. The stimulus-response formula which leads to respondent behavior cannot explain however a large part of behavior which is not under the control of eliciting and obviously recognizable stimuli and therefore is termed *operant behavior*

> An operant then represents an identifiable part of behavior for which no correlated stimulus needs to be detectable upon occasions when it is observed to occur It is studied as an event appearing spontaneously with a given frequency.

Since the term reflex applies to both operant and respondent behavior, a reflex is recognized by each action of the lever and by relating the number of actions to time as represented in curves, from which properties of the reflex can be recognized and estimated from the slope. Inducing an animal to perform the task mentioned is rather simple. It is therefore possible to introduce into the chain of events a number of modifying factors which are useful for further analysis After a description of the experimental method employed (Chapter II) the following chapters, III to VIII, give an account of Skinner's experimental data. It is shown how the action of the rat can be reinforced, as evident in an increase of the slope of the performance curves, by interjection of the release of a pellet by the observer, and how delay of reinforcement affects the time course of the building up of responsiveness. Recovery from extinction by reinforcement and spontaneous recovery are discussed Inhibition, negative conditioning, and abolishment of a response are demonstrated. Time factors in relation to extinction and reinforcement, and periodic reconditioning, are shown to play an important rôle in the establishment of responses. Furthermore the differentiation of a response is investigated in relation to extinction after a differentiation has been established.

The experimental results are illustrated by graphs which help greatly in following the diversity of various experimental tests employed. The section containing the experimental data undoubtedly would have gained considerably in clarity if a summary of results had been added to each chapter. For simplification a system of

Notation is introduced which is used throughout the course of the book and which states for each chapter the relations of stimulus and response in building up the behavioral picture Chapters IX and X, dealing with drive and its influence upon animal performance, are of interest as indicating how deprivation, thirst, running, emotions, drugs, and diurnal rhythms of spontaneous activity affect behavior and could for any experimenter contain a number of good suggestions if the known facts were sufficiently stressed Chapter XII presents an analysis of the behavior in relation to properties of the nervous system as deduced from Skinner's experimental results To summarize Skinner's conception the following sentences may be quoted.

> A quantitative science of behavior may be regarded as a sort of thermodynamics of the nervous system It provides *descriptions of the activity of the nervous system of the greatest possible generality* Neurology cannot prove these laws wrong if they are valid at the level of behavior. Not only are laws of behavior independent of neurological support, they actually impose certain limiting conditions upon any science which undertakes to study the internal economy of the organism The contribution that a science of behavior makes *to neurology is a rigorous and quantitative statement of the* program before it

We have to ask ourselves, however, whether anyone ever presumed to find out anything about the function of the nervous system except by analysis of behavior of single end organs, fibers, or a whole organism?

Chapter XIII *contains a conclusion, in which objections* are raised against "systems" of behavior such as Lewin's, Hull's, and the concept of tropisms as satisfactory explanations for behavior. In tropistic experiments

> variables have been isolated which are capable of being treated quantitatively and which behave in lawful ways. They *may also* be combined in larger complexes with predictable effects But any system which takes orientation or oriented progression as the only property of behavior to be accounted for and which requires a stimulus only as a field of force is seriously circumscribed In the case of higher organisms at least it is presumably possible to set up an independent

> descriptive system based upon the concept of the reflex that will yield an equally satisfactory result. Where behavior is largely orientation and where stimuli are fields of force we may prefer the concept of the tropism on grounds of simplicity while at the same time rejecting it in the case of more complicated organisms.

What Skinner rejects here as methodically and for explanatory reasons inadequate for study of behavior, he has already accepted, however, as a convenient tool for the experimental studies presented in his own book (Chapter I, p. 6) and he uses it throughout. Furthermore it needs to be emphasized that the biologists he criticizes have not proposed to build up any "system of behavior" in the sense of Skinner's system of behavior on the basis of tropisms alone, and indeed have specifically denied this, they have rather collected data which for the convenience and predictive value can somehow be classified systematically. The method of definition for functional properties as presented by Skinner is not new.

The book as a whole is undoubtedly too exclusively personal. This refers as well to the language used as to the criticism applied to the literature of the field. It is certainly striking not to find Lashley nor any of his outstanding work mentioned. There is only one reference to an eleven-year-old article by Richter, to mention only two cases which indicate that the promising title of the book does not live up to the expectation of the reader. Experiments with the albino rat are certainly not enough to establish a system of behavior. There are many other animals which might be successfully employed for experimentation as well and help to clarify significantly our concept of behavior. Furthermore the book does not give any demonstration of behavioral invariants, since only one strain of rats has been used for tests, whereas it is well known from numerous studies that by peculiarities of different strains it is possible to obtain such information. It therefore seems doubtful whether a system of behavior can be of any value as long as its applicability has not stood the test with numerous other animals besides the white rat. Skinner really ought to furnish proof of this before elaborating a *system*. It has to be said, however, that the experimental data as they stand, and the simplicity of experimental procedure, are certainly of significant value and vastly better for analytical purposes than learning curves and the like. From the standpoint of the reviewer as a physiologist the title of the book promises a lot, but after having read it one is inclined to propose for it a title which would rather read, "Properties of certain specific complex reflex mechanisms in the white rat"

The Biological Laboratories
Harvard University
Cambridge, Massachusetts

[*Spearman, C Psychology Down the Ages (Two vols) New York Macmillan, 1938. Pp. 454 and 355*]

REVIEWED BY HERBERT MOORE

One may write a history of psychology in the form of a chronological series of events, discoveries, or important men, many of the older histories conform to this pattern. Another historian may be concerned with the development of ideas, without any regard to their chronological place, Gardner Murphy's, Muller-Freienfels', and Spearman's follow this pattern Each one of these latter three is biased by the pet psychological system, and limited by the interests, of the author, Murphy's fails to discuss many vital issues in the development of psychology, Freienfels' is unduly concerned with continental interpretations and theoretical implications, and Spearman's is written in such a manner that it seems to hope to leave the reader with the conviction that the history of psychology has been a propaedeutic to the formation of three noegenetic laws. Of these three histories, Spearman's is undoubtedly the most erudite. Most of the concepts which have concerned philosophers and psychologists since thoughts were recorded are evaluated in the light of "how far and in what direction they have gone beyond the confines of common sense" (I, p 6), and reviewed to discover the manner in which "the bold attempts to bring the constituents of experience under the rule of laws" (I, p 7) have either consistency or validity. Having formulated such a plan, Spearman might be expected to enunciate the laws of behavior, (*a*) in a general sense as they apply to total organic reactions, and (*b*) in a specific sense as they affect or interpret the modes of behavior of comparatively separable sense or functional entities, instead he contents himself with an appraisal and exposition of those concept-forming laws which for the past fifteen years have been identified with his name, and tends to evaluate other methods and contributions in terms of their harmony with the principles on which these laws are founded The end result is an undue emphasis on the value of statistical findings and theoretical implications, and a nigh total disregard for such experimental support as has been given to some theories and laws as to warrant their adoption as established facts in any exposition of the modern systems of psychology

In spite of the shortcomings which such a limitation necessitates, Spearman has provided the world of psychology with a much-needed perspective, a need which could only be filled by one as familiar with the history of thought as Spearman obviously is His two volumes are by no means a superficial history of philosophical psychology; his work is never second hand, in fact he rarely quotes the most reliable historians Authorities, be they Greek or Roman, Oriental or Arabian, are the direct source of his excerpts, and he weaves them together into segregated thought-patterns, in which Aristotle, Aquinas, Hierocles, Mill, Husserl, and many other familiar (and unfamiliar) authors are repeated page-mates; the result is a contribution which only those familiar with the history of philosophy and psychology can appreciate, one which could well complement James' *Principles,* and supply the raw data for advanced courses in the history of psychology.

The two volumes are in five sections. What psychology is about, (91 pp), what the psyche can do (89 pp.), how the psyche is constituted, (257 pp) what follows what (167 pp), and what goes with what (115 pp.) Forty pages of notes which quite often give valuable references to, and comments on, casually mentioned topics and problems accentuate its value as a reference book. Discreteness rather than continuity characterizes the discussion of each subtopic, quite often the ideational antecedents of major psychological concepts are given as interrupted ideas in one subtopic and not even mentioned when the concept is being evaluated, e.g , Chapter 17, pp. 287-317, including a discussion of impulse, will, and purpose, is directly related to the concept of force which is discussed on pp. 66-67, Chapter 13, pp 214-236, the perception of relations, and Chapter 24, pp 419-446, the confusion that is gestalt psychology, are directly connected with the concept of form which is introduced on pp 70-72 and 95-99 It is this jumping around from topic to topic and snatching kernels from different thought samples that militates against the adequacy of the book as a source for a thorough discussion of any phase of psychology except noegenetic laws. Even as a concept-evaluative treatise, the constant tendency to make appraisals in terms of common sense rather than in terms of experimental verifications destroys its validity for the scientist. Throughout each section and subtopic, Spearman *interjects his own preferences* and point of view, and never fails to leave the reader without a pronunciamento on the value of the *modern version of the concept.*

Section I defines the province of psychology, suggests psyche as a more desirable subject *for appraisal than mind or behavior, admits* the value of physiology only to the extent that it provides tentative hypotheses for explaining mental activity, appraises such psychological methods as self-knowledge, intuition, induction, analogy, mathematics and experiment, and sums up the value of psychological issues of importance to the psychologist with the statement

> *Psychological discussion is now being obscured and its* progress arrested by the introduction of these philosophical issues that seemingly cannot be brought to any decision . nevertheless, these same topics are to a certain extent indispensable, completely deprived of them psychology would become ridiculous (I, p *76*)

Section 2 surveys the development of the faculties in psychology, lists the common objections against them and the apparent result of their abandonment, anarchic chaos, and suggests that a solution of the dilemma in which psychology finds itself can be found only when the constituents of the psyche are analyzed and synthesized, and their functional interrelations explained Basic psychological terms are evaluated Intelligence has excited the greatest enthusiasm and to it "the greatest amount of labor has been devoted" On its nature every authority differs, and that difference is all the more significant because of the claim that its degree can be measured by tests, the most of which are "trifling stunts" Attention still retains the Platonic concept *of intellectuality, but that has been* sprinkled over with the Stoic concept of tension, thus producing the *modern interpretation, strained intellect.* Sensation still retains its two-fold interpretation, the eyes and ears of the soul, and the *antithesis of intellect* Memory, a late arrival, has been given sufficient definiteness for scientific isolation and discussion, while imagination is still in the hazy state of indefiniteness The orectic faculties have always been given a dynamic function, and have been described in the same manner, though with different names, in the course of history

Section 3 surveys "the basis of sensory perception" with a review and evaluation of the four common elements, the relation of sensory stimuli to sensory percepts, the relation between the sensory perception of material objects and the objects as conceived by physical science, and the extent to which sensory perceptions are amenable to analysis; the "perception of relations" with a review of the significance that has been given to scheme, figure, form, discrimination and comparison in psychological literature since Plato's day, and the suggestion that "relations and forms are really terms denoting the very same objective facts perceived, but indicating more or less different manners of perceiving them" (I, p 236), "perceptual supplements" with a résumé of the suggestions that have been offered to account for meanings that are given to sensory experiences—reproducts and supersensibles—but with very little attention to the experimental battles that have been waged around the place of images, "ways of regarding"; "mental states", "complexes of behavior", "unconscious mind", "the I and the self", and finally a chapter on "the confusion that is Gestalt psychology" which will probably provoke more controversy than any other chapter in the two volumes The various interpretations that have been given Gestalt—form, sensory appearance, underlying force, and wholeness—are reviewed and evaluated, and from the review the inference is drawn that "there seems no escape from the conclusion that there is something somewhere rotten in the state of Gestalt psychology" (I, p. 436), the crux of which rottenness lies in the varied and confused interpretations given the term "wholeness," their failure to distinguish between subjective and objective experiences, their assumption of the temporal and functional priority of wholes to parts, and in their separation of passive perceptual experiences from active perceptions, thoughts, feelings and volitions

Section 4 climaxes the two volumes. For its sake the preceding three sections are written. It is primarily concerned with an evaluation and reformulation of psychological laws. After setting up a standard in terms of which all laws may be appraised, the laws that have been offered at various times in history are judged. The standard set up as a criterion for a psychological law is,

> that they should afford power of foreseeing the future as well as divining the past In order to secure this power the most important conditions have shown themselves to be in the first

> place, that the course of events should proceed uniformly and be amenable to analysis, in the second place, that the laws found to govern the events should have the greatest possible generality and ultimacy (II, p. 20)

Judged by these standards, the law of association is "an enormous advance upon anything within the confines of common sense"; the law of dispositions has been established and become "one of the revolutionary forces of society" in its application to problems of memorizing and the transfer of training, but has had very little use in the fields of character development and disintegration, the law of retentivity, that "every mental process lags behind the stimulus in beginning and much more so in stopping," implies two phases of one basic process; the law of control, that "the amount of cognition may be directly controlled by conation," stands on a much inferior footing to that of association and has received from scientific psychology very little more than "damaging exaggerations", the law of constant output, that "every mind tends to keep its total simultaneous output constant in quantity, however varying in quality," has all the characteristics of a scientific law, but is lamentably deficient in verification and study, the law of fatigue is "among the greatest conquests in psychology" and in large measure satisfies the essential requirements of science—foresight and fundamentality

Spearman supplements his survey of these "commonly established laws" with "a very ambitious adventure . . . a complete set of laws expressing and explaining the origin and evidence of all knowledge" (II, p. 142)—his three noegenetic laws, and the attempts made at the formulation of laws in the spheres of impulse, will, output, and control, thus he provides a scheme of four groups of laws, (*a*) primitive states, (*b*) cognitive quality—knowledge of experience, knowledge of relations, and conception of correlates, (*c*) orectic quality—impulse and will, and (*d*) cognitive and orectic quantity—output, control, retentivity, fatigue and basal conditions

Section 5 is in the nature of an appendix correlational coefficients supply the demand that was expected to be met by temperaments and types, the factor theory is a de novo creation and has made steady and marked progress since its inception, the tetrad method provides more definite conclusions than those suggested by Thurstone and Kelley, and W is as basic to orexis as G is to cognition.

Whether or not the importance that Spearman gives to the *enunciation of* laws is warranted depends upon the bias of the interpreter. At their best laws are conclusions resulting from the collection and interpretation of large bodies of data. Their enunciation could, by no stretch of the imagination, be called a history of a science. Their establishment may be milestones in the development of *a fund of knowledge, when there is common agreement regarding* their validity. It is scarcely worthwhile pointing out that many *of those laws which Spearman accepts as established have not gone* through much of a grist of experimentation, and are still in the realm of armchair speculation.

As *a survey of the contribution of psychology to the world of* knowledge, the book can lay no claim to either adequacy or comprehensiveness. Dubbing systems with epithets is neither description nor evaluation. To describe associationists as those "whose handling of delicate matters is not seldom suggestive of bovine activity in a china shop" (I, p. 23), introspectionists as those whose "ambitious psychology soon underwent an almost unparalleled collapse. As it rose like a rocket it fell like a stick" (I, p. 81), behaviorism as "suicidal psychology which is hard to assimilate . . . the fate of which baffles the imagination" (I, pp. 310-311), and Gestalt psychology as "psychological enemy No. 2" (I, p. 446)—is to offer neither an explanation nor a condemnation of the systems. An appraisal of any one system in terms of the assumed laws of another system is foredoomed to disaster.

Viewed as a background of Psychology à la Spearman, the two volumes will undoubtedly remain as landmarks in the evolution of modern psychology, viewed as "psychology down the ages" they do not provide a comprehensive picture of the development of psychological *concepts or systems, and leave the sophisticated with the impression* that there has been undue emphasis on the evaluative and utilitarian aspects at the expense of unbiased appraisal of the various suggestions that history provides in the various phases, speculative and experimental, of the province of psychology.

Department of Psychology
Mount Holyoke College
South Hadley, Massachusetts

BOOKS RECENTLY RECEIVED

(There will always be two pages of book titles, listed in the order of receipt, i e, the most recently received books will be found at the end of the list)

McFARLAND, M B Relationships between Young Sisters as Revealed in Their Overt Responses New York Bur Pub Teach Coll, Columbia Univ, 1938 Pp 230

BECK, S J Personality Structure in Schizophrenia New York Nerv. & Men Dis Monog, 1938, No 63 Pp 88

RUDOLFER, N DA SILVEIRA An Introduction to Educational Psychology. Sao Paulo Companhia Editora Nacional, 1938 Pp 459

JACOBSON, E You Can Sleep Well New York McGraw-Hill, 1938. Pp. 269

LANDTMAN, G The Origin of the Inequality of the Social Classes Chicago Univ Chicago Press, 1938 Pp 444

SIMPSON, R H A Study of Those Who Influence and of Those Who Are Influenced in Discussion New York Bur Pub Teach Coll, Columbia Univ, 1938 Pp 89

BENDER, L A Visual Motor Gestalt Test and Its Clinical Use New York Amer Orthopsychiat Assoc, 1938 Pp 176

KIRKPATRICK, C Nazi Germany Its Women and Family Life New York, Bobbs-Merrill 1938 Pp 353

MURRAY, H A Explorations in Personality New York, Oxford Univ Press, 1938 Pp 761

[illegible], A C Representative Psychologists Union Bridge, Md Pilot Publishing, 1938 Pp 119

[illegible], G La Psychiatrie et les Sciences de l'Homme Paris Alcan, 1938 Pp 194

HODGSON, S The Man Who Made the Peace: Neville Chamberlain New York Dutton, 1938 Pp 147

SPENCER, D Fulcra of Conflict New York World Book, 1938 Pp 306

LINK, H C The Rediscovery of Man New York Macmillan, 1938 Pp 257

[illegible], M M The Inter-Trait Rating Technique New York Teach. Coll, Columbia Univ, 1938 Pp. 99

FRAMPTON, M E, & ROWELL, H G Education of the Handicapped Vol 1 New York World Book, 1938 Pp 260

ABEL, J J Chemistry in Relation to Biology and Medicine Baltimore Williams & Wilkins, 1938 Pp 79

FIREMAN, P Life Without Fear New York Dutton, 1930 Pp. 111

BURLOUD, A Principles d'une Psychologie des Tendances Paris Alcan, 1938 Pp 427

DUNBAR, H F Emotions and Bodily Changes New York Columbia Univ Press, 1938 Pp 601

ROBERTS, H The Troubled Mind New York Dutton, 1939. Pp 284

NAGEL, E Principles of the Theory of Probability Chicago Univ Chicago Press, 1939 Pp 80

MOORE, H Psychology for Business and Industry New York McGraw-Hill 1939 Pp 527

SINNOTT, E W, & DUNN, L C Principles of Genetics New York McGraw Hill, 1939 Pp 408

RASHEVSKY, N Mathematical Biophysics Chicago Univ. Chicago Press, 1938 Pp 340

JERSILD, A T, & FITE, MARY D The Influence of Nursery School Experience on Children's Social Adjustments Bur Pub, Teach Coll, Columbia Univ, 1939 Pp 112

RICKMAN, J. A General Selection from the Works of Sigmund Freud. *London: Hogarth Press, 1937. Pp. 329.*

KLEIN, M., & RIVIERE, J. Love, Hate, and Reparation. London: Hogarth Press, 1937. Pp. 119.

POWDERMAKER, HORTENSE. After Freedom. New York: Viking, *1939.* Pp. 408.

REED, H. B. Psychology and Teaching of Secondary School Subjects. New York: Prentice-Hall, 1939. Pp. 684.

ROGERS, C. R. The Clinical Treatment of the Problem Child. New York: *Houghton Mifflin,* 1939. Pp. 393.

GESELL, A., CASTNER, B. M., THOMPSON, H., & AMATRUDA, C. S. Biographies of Child Development. New York: Harper, 1939. Pp. 328.

DUNLAP, J. W. Workbook in Statistical Method. New York: Prentice-Hall, 1939. Pp. 145.

KUHN, E. G. The Pronunciation of Vowel Sounds. Bur. Pub., Teach. Coll., Columbia Univ., 1938. Pp. 86.

KOHLER, W. The Place of Value in a World of Facts. New York: *Liveright, 1938.* Pp. 418.

DOLLARD, J., DOOB, L. W., MILLER, N. E., MOWRER, O. H., SEARS, R. R., *et al.* Frustration and Aggression. New Haven: Yale Univ. Press, 1939. Pp. 209.

PRATT, C. C. The Logic of Modern Psychology. New York: Macmillan, 1939. Pp. 185.

LYND, R. S. Knowledge for What? Princeton Univ. Press, 1939. Pp. 268.

LUND, F. H. Emotions: Their Psychological, Physiological, and Educative *Implications. New York: Ronald Press, 1939. Pp. 305.*

BEACH, W. G. The Growth of Social Thought. New York: Scribner's, 1939. Pp. 250.

GILLILAND, A. L., & CLARK, E. L. Psychology of Individual Differences. New York: Prentice-Hall, 1939. Pp. 535.

MONEY-KYRLE, R. Superstition and Society. London: Hogarth Press, 1939. Pp. 163.

RICKMAN, J. Civilization, War, and Death. Selections from Three Works *by Sigmund Freud. London: Hogarth Press, 1939. Pp. 102.*

WHEELER, W. M. Essays in Philosophical Biology. (Ed. by G. H. Parker.) Cambridge: Harvard Univ. Press, 1939. Pp. 261.

WINN, R. B. Scientific Hypnotism. Boston: Christopher, 1939. Pp. 168.

THORNTON, H., & THORNTON, F. How to Achieve Sex Happiness in Marriage. New York: Vanguard Press, 1939. Pp. 155.

FREEMAN, E. Principles of General Psychology. New York: Holt, 1939. Pp. 530.

SCHEIDEMANN, NORMA V. *Lecture Demonstrations for General Psychology.* Chicago: Univ. Chicago Press, 1939. Pp. 241.

LANDIS, C., & HUNT, W. A. The Startle Pattern. New York: Farrar & Rinehart, 1939. Pp. 168.

GUTHEIL, E. A. The Language of the Dream. New York: Macmillan, 1939. Pp. 286.

KENDALL, G. M. Organization and Teaching of Social and Economic Studies in Correctional Institutions. Bur. Pub., Teach. Coll., Columbia *Univ., 1939.*

LOWENFELD, V. The Nature of Creative Activity. New York: Harcourt Brace, 1939. Pp. 272.

MIKESELL, W. H. Mental Hygiene. New York: Prentice-Hall, 1939. Pp. 456.

ROSETT, J. The Mechanism of Thought, Imagery, and Hallucination. New York: Columbia Univ. Press, 1939. Pp. 289.

Genetic Psychology Monographs (*continued*)

VOLUME 17—January-December, 1935

1. Organization of Behavior in the Albino Rat—R. L. Thorndike
2. Brightness Discrimination in the Rhesus Monkey—M. [illegible] Crawford
3. The Limits of Learning Ability in Cebus Monkeys—A. M. Koch
4. [illegible]—A. H. [illegible]
5. On Intelligence of Epileptic Children—E. B. Sullivan and L. Gahagan
6. A Study of the Play of Preschool Children by an Unobserved Observer—[illegible]

VOLUME 18—January-December, 1936

1. Sex Differences in Variational Tendency—Q. McNemar and L. M. Terman
2. The Process of Learning to Dress among Nursery-School Children—C. B. Key, M. R. White, W. P. Honzik, A. B. Heiney, and D. Erwin
3. A Study of the Present Social Status of a Group of Adults Who, When They Were in the Elementary Schools, Were Classified as Mentally Deficient—Warren R. Baller
4. The Influence of Specific Experience upon Mental Organization—Anne Anastasi
5 & 6. Studies in Aggressiveness—Lauretta Bender, Sylvan Keiser, and Paul Schilder

VOLUME 19—January-December, 1937

1. Psychological Bases of Self-Mutilation—Casimir Dabrowski
   [illegible] A Study of the Relationship Between [illegible] and [illegible]—[illegible]
2. A Psychological Study of Forty Unmarried Mothers—Ruth D. Nottingham
   Behavior Problems in the Children of Psychotic and Criminal Parents—Lauretta Bender
3. Dominative and Integrative Behavior of Young Children in an Experimental Play Situation—Harold H. Anderson
4. The Sequential Patterning of Prone Progression in the Human Infant—Louise Bates Ames

VOLUME 20—January-December, 1938

1. The Relationship Between Characteristics of Personality and Physique in Adolescents—P. S. de Q. Cabot
2. Behavior Problems of Elementary School Children: A Descriptive and Comparative Study—[illegible]
   Graphic Representation of a Man by Four-Year-Old Children in Nine Prescribed Drawing Situations—[illegible]
3. Differences Between Two Groups of Adult Criminals—[illegible]
4. A Comparative Study by Means of the Rorschach Method of Personality Development in Twenty Pairs of Identical Twins—Evelyn Troup
   Individual Differences in the Facial Expressive Behavior of Preschool Children: A Study by the Time-Sampling Method—[illegible]

VOLUME 21—January-December, 1939

1. An Experimental Analysis of "Level of Aspiration"—Rosalind Gould
   [illegible]
2. Some Light on the Problem of [illegible] as Found from a Study of [illegible]
   Preschool Children of Non-American Ancestry [illegible]
3. Domination and Social Integration in the Behavior of Kindergarten Children and Teachers—[illegible]
   The Capacity of the Rhesus and Cebus Monkey and the Gibbon to Acquire Differential Response to Complex Visual Stimuli—[illegible]

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.